SECURITY CLASSIFICATION OF THIS PA

RE

AD-A244 658

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION | MARKINGS

2a. SECURITY CLASSIFICATION AUTHOR | 3. DISTRIBUTION/AVAILABILITY OF REPORT

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE

ELECTE JAN 13 1992 S D

Approved for public release; distribution unlimited.

(2)

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)

AFOSR-TR- 92 0027

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
The Smith-Kettlewell Eye Research Institute		Air Force Office of Scientific Research/NL

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
2232 Webster Street San Francisco, CA 94115	Bldg. 410 Bolling AFB, D.C. 20332-6448 Attn: John F. Tangney

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR	NL	FQ8671-9001374 AFOSR-89-0035

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
Building 410 Bolling AFB DC 20332-6445	61102F	2313	A9	

11. TITLE (Include Security Classification)
"Visual Processing of Object Velocity and Acceleration"

12. PERSONAL AUTHOR(S)
McKee, Suzanne

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Technical Report	FROM 88/10/15 TO 91/10/14	91/12/13	7

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Six separate projects have explored how velocity and acceleration are encoded in the human visual system. 1) Welch demonstrated speed discrimination for coherent plaid patterns formed of two superimposed gratings was limited by the speed of the gratings, not the apparent speed of the plaid itself. 2) Bowne, and more recently Grzywacz, applied "motion-energy" models to the psychophysics of speed discrimination. 3) McKee and Welch compared the relative precision of velocity and size constancy, finding little evidence for velocity constancy in human motion processing. 4) Watamaniuk demonstrated that the visual system integrates diverse speeds (2 - 8 deg/sec) in a random dot display to obtain a precise estimate of the mean speed. 5) McKee and Watamaniuk found that a single point (the "signal") moving in apparent motion along a fixed trajectory was easily detected amidst other similar points in random apparent motion (the "noise"), even though the spatial and temporal characteristics of the signal and noise points were identical on a frame-by-frame-basis. 6) Bravo and Watamaniuk showed that two sets of randomly distributed dots moving in the same direction, but at two very different speeds, formed two transparent planes; discrimination of small changes in the speed of one set of dots was unaffected by the presence of the other dots.

92-01027

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☑ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	(U)

22a NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Dr John F Tangney	(202) 767-5021	NL

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

16 DEC 1991

92 1 10 010

SUMMARY

The work from this laboratory for the past three years can be divided into six areas: 1) Dr. Welch's studies on motion coherence and transparency in plaid patterns; 2) Dr. Bowne's and, more recently, Dr. Grzywacz's applications of filtering models to the psychophysics of speed discrimination; 3) the McKee-Welch studies on the precision of the visual constancies; 4) Dr. Watamaniuk's study on the integration of speed information in globally-defined motion; 5) the McKee-Watamaniuk studies on local motion trajectories in the midst of random motion; 6) the Bravo-Watamaniuk studies on motion transparency induced by speed differences in random dot displays.

A. The Welch Results

The direction and speed of a long moving line or a grating viewed through a circular aperture is ambiguous because only the motion perpendicular to the orientation of the line or grating can be detected. In a pattern composed of several oriented contours, this ambiguity may be resolved either 1) by assigning the unambiguous motion of small local features, such as the ends of lines, the corners, or the contour intersections, to the pattern as a whole, or 2) by appropriately combining the motion vectors from each contour to identify their common velocity. Consider a plaid pattern formed of two superimposed moving gratings. The plaid appears to move with the velocity of the nodes formed at the positions where the two gratings intersect.. Is this a case where the whole pattern has been assigned the unambiguous motion of these local nodal features or has the motion of each grating been assessed separately and then combined? Several years ago, Adelson and Movshon (1982) proposed a two-stage solution to the aperture problem, using the motion of a plaid as an example of this two-stage motion processing. They suggested that initially the moving plaid was decomposed into the one-dimensional motion components associated with each of the gratings, and then the components were recombined according to the intersection of the two constraint lines associated with each grating. The motion of each grating is consistent with a family of possible velocity vectors -- all the velocity vectors that are constrained to have the same motion component on the axis perpendicular to the orientation of the grating (See Figure 1). The intersection of the two lines of constraint for the two gratings corresponds to the velocity of the nodal points. Adelson and Movshon suggested that the perceived motion of the plaid depended on a neural representation of this intersection of constraints occurring at a stage that followed the initial processing of the grating components, and they presented convincing psychophysical data supporting this two-stage model. Nevertheless, it was still possible that the human observer was really assigning the directly-perceived velocity of the nodal features to the plaid, without the intervention of two processing stages.

Dr. Welch provided compelling psychophysical evidence for the two-stage model. In a paper published in Nature, she showed that thresholds for discriminating the speed of the plaid depend on the speeds of the gratings forming the plaid, not the plaid itself. The plaid necessarily moves faster than the gratings that form it. In her study, Dr. Welch used a plaid pattern that moved five times faster than its grating components. Speed discrimination for very slow speeds (< 1 deg/sec) is not very precise. Dr. Welch was able to show that the discrimination of the plaid speed was much less precise than for a one-dimensional grating moving at a speed equivalent to the speed of the plaid (or the nodal points), because speed discrimination for the plaid was,

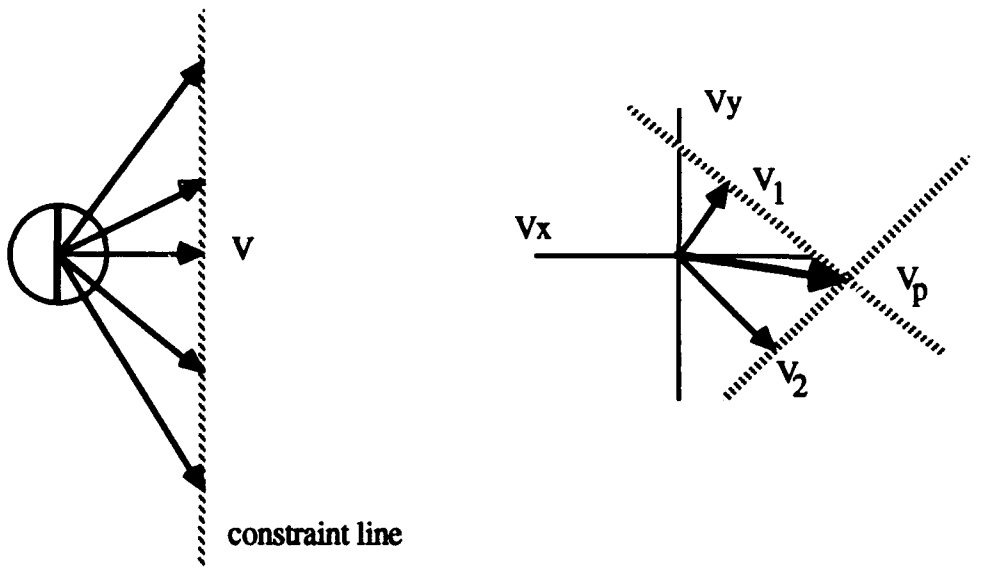


Figure 1

in fact, limited by the much slower speed of the gratings forming the plaid. Although speed discrimination for the plaid is limited by noise at the site where the speed of the gratings is encoded, Dr. Welch demonstrated that the human observer does *not* have access to this information if the two gratings cohere into a single moving pattern. To show the inaccessibility of the grating speed, she first established that observers could easily judge grating speed even if the orientation of the grating changed randomly from trial-to-trial. She then superimposed a pair of gratings, each with random variations in orientation, and found that observers could no longer discriminate the speed of the resulting plaid pattern, because the perceived speed of the plaid necessarily changed randomly from trial-to-trial since the angle between the gratings affects plaid speed.

When two moving gratings are superimposed, they do not always cohere. They sometimes slide over one another like two transparent patterns. Simply stated, coherence depends on the similarity of the two gratings; the more similar the gratings, the more likely they are to cohere. In a second paper (Perception, 1991) Welch and Bowne demonstrated that perceived coherence is an indicator of whether the observer has access to the low-level signal produced by the gratings, or only has access to signals from the second stage of motion processing after the signals from the gratings have been combined. They found that the speed of the gratings was inaccessible when the gratings cohered but that observers could easily make judgments about the speed of the gratings when the gratings appeared transparent. In her dissertation work, Dr. Welch used the speed discrimination paradigm to explore what is meant by "similar" in the context of human motion processing. Gratings cohere when their contrast and spatial frequency are similar. However, it is similarity of speed, not temporal frequency that determines coherence for these plaid stimuli. The coherence rules also operate to determine whether thin line targets (and presumably other contours) cohere. For the talk presented at the Optical Society Meeting in 1991, Dr. Welch brought some demonstrations of these effects for line targets.

B.The Bowne-Grzywacz Work

Can the precision of speed discrimination be explained by "motion energy" detectors of the type proposed by Adelson and Bergen to explain direction discrimination? Is speed discrimination limited by noise in the spatial and temporal filtering that occurs early in the visual pathways? Dr. Bowne's careful calculations indicate that "motion energy" detectors (or "Reichardt correlators") cannot account for our speed discrimination results. We used a very elementary stimulus configuration to examine the discrimination of temporal signals -- speed discrimination for two-frame apparent motion. The stimulus consisted of a four-frame sequence of bright points hopping in apparent motion. There was a fixed asynchrony between the outer pair of points (the first and last), but the time between the inner pair of points varied from trial-to-trial. The task of the observer was to discriminate the time (speed) between the inner pair of points. In a paper published in the Journal of the Optical Society of America (1989), Bowne, McKee and Glaser showed that the outer pair of points significantly degraded the ability to judge the speed of the inner points even when they were separated by distances as large as 1 degree or by temporal intervals as large as 200 msec. Small-scale (high spatial frequency) "motion energy" units should have easily detected the local signal arising from the inner pair of points without detecting the signal produced by the outer interfering pair. Thus, one might conclude that speed is mediated only by large-scale (low -frequency) motion energy units. However, speed discrimination for the inner test pair was significantly degraded even when the outer interfering targets were high spatial-frequency "Difference of Gaussians" targets of low contrast, targets that would be nearly invisible to the large-scale units. Dr. Bowne concluded that speed discrimination depends on interactions between signals from many different scales operating over large distances, a conclusion that would support current velocity models such the one proposed by Heeger (1987) and also by Grzywacz and Yuille (1990).

In a second study, Bowne made detailed measurements of the sensitivity of human temporal mechanisms, using the two-pulse subthreshold summation paradigm (Rashbass, 1970; Watson and Nachmias, 1977) with drifting sinusoidal targets. He also measured the contrast sensitivity and contrast discrimination functions for drifting gratings. Dr. Bowne then attempted to model the speed discrimination thresholds for these same drifting sinusoidal targets using the line-element approach (Wilson, 1986) and/or "viewprint" approach (Klein and Levi, 1985) that have worked in modeling human discrimination in hyperacuity judgments. Surprisingly, speed discrimination at low contrasts was more precise than could be predicted from any combination of the temporal mechanisms, although our estimates of human temporal mechanisms did an excellent job of predicting *local* temporal discrimination, i.e., the ability to discriminate between slow and rapid onsets. Again these results suggest that speed discrimination is accomplished by summing signals over large spatial areas. It is interesting that recent physiological measurements of the contrast sensitivity of neurons in cortical area MT indicate that some of these large velocity-tuned units are more sensitive than neurons encountered at earlier stages of visual processing (Sclar, Maunsell and Lennie, 1990).

In 1989, Grzywacz and Yuille published a model of visual velocity computation that employed a population of spatio-temporally oriented filters to encode velocity. Dr. Grzywacz has attempted to reconcile his model with a variety of psychophysical observations that appear to challenge this approach. By introducing a rectified band-pass filter in front of the motion-energy filters (Chubb & Sperling, 1988), Dr. Grzywacz was able to explain "non-Fourier" motion,

"beat-pattern" motion and the invariance of speed discrimination with increasing contrast.

C. The McKee-Welch Studies of Constancy

In traditional studies of size constancy, observers were often shown an object at some faraway distance and were asked to adjust the size of an adjacent object until it matched the distant object (Holway and Boring, 1941). Sometimes, the aim of these studies was to determine what was actually seen by the observer -- the objective size or the angular size?. In other cases, the intent was to explore limitations on size constancy, e.g., over what distances could observers match objective size before perhaps regressing to match based on angular size. In one of the most interesting of these studies, Gilinsky (1955) found that observers were able to match either the retinal or the objective size of the test object, depending on the instructions given by the experimenter. Gilinsky's results, subsequently verified in other laboratories (Carlson, 1960; 1977; Leibowitz and Harvey, 1969), indicate that matching is a weak guide to the cognitive (or neural) operations underlying size constancy. It is not possible to determine whether the observer perceives retinal size, and then corrects this percept by some measure of depth to estimate objective size, or *vice versa.*

There is a psychophysical tool that could reveal the coding sequence. Instead of asking observers what they perceive, the relative *precision* of their judgments of angular and objective size can be determined. What is the smallest detectable change in objective size? What is the threshold for discriminating differences in angular size? The precision of psychophysical thresholds is limited either by noise in the stimulus itself, or by noise in the neural pathways coding the stimulus dimensions -- more noise means less precision. If the calculation of objective size involves the simple combination of two independent neural measurements (angular size and depth), objective size judgments should be consistently less precise than angular size judgments, because the depth measurement will add noise to the calculation. Burbeck (1987) used this approach to measure spatial frequency discrimination for low spatial frequency targets, and found surprisingly, that judgments of objective spatial frequency (in cyl/cm) were as precise or perhaps slightly more precise than judgments of angular spatial frequency (in cyl/deg). Her results indicate that human observers do not have access to information about angular size. In a variant of Burbeck's study, subjects were asked to judge small changes in the lateral distance separating a pair of lines, while target disparity was randomly varied over a $\pm$ 40 arc minute range from trial-to-trial (paper accepted by Vision Research). We confirmed Burbeck's results for line targets separated by large lateral distances, but for small distances ($\leq$ 20 arc minutes), angular size discrimination (minutes of arc) was superior to objective size discrimination (cm).

Our results also indicate that size constancy may be learned. We asked our subjects to make judgments of objective size (cm) when the angular size of the target was manipulated so that it increased with increasing three-dimensional distance (defined by disparity in a stereoscope), rather than decreased as occurs in natural circumstances. In this "anti-constancy" condition, subjects were able to make fairly good judgments of objective size after only a small amount of practice (600 trials). Their judgments were less precise than in the condition simulating natural constancy, but perhaps with more practice, subjects could become skilled at "anti-constancy" as well.

In an earlier study (Vision Research), McKee and Welch used measures of precision to compare velocity constancy to size constancy under identical conditions. They found that observers were unable to use disparity information to transform the angular velocity signal into a

precise object-based code. The Weber fraction for discriminating changes in objective velocity (cm/sec) was about twice the Weber fraction for discriminating changes in angular velocity (deg/sec), and was substantially higher than predicted from a combination of the errors in judging disparity and angular velocity. By comparison, judgments of the distance traversed by the moving target showed excellent size constancy. The discrimination of changes in objective size (cm) was as precise as the discrimination of changes in angular size (deg). The angular velocity signal is useful without transformation into an object-based signal; it guides eye and body movements, and is the basis of motion parallax judgments. The need to retain this angular signal may explain why there is no efficient mechanism for velocity constancy.

D. The Watamaniuk Speed Discrimination Study

Williams and Sekuler (1984) demonstrated that a motion stimulus composed of many spatially-intermingled motion vectors chosen at random from a broad range of directions (bandwidth 180 degrees or less) produced a percept of global flow, moving in a direction approximately equal to the mean of the range. In work begun in Dr. Sekuler's laboratory, Dr. Watamaniuk measured speed discrimination for a stimulus composed of many spatially-intermingled motion vectors that all moved in the same direction, but with a wide range of different speeds. He found that subjects could easily discriminate between the *mean* speeds of these distributions (for some conditions the velocity Weber fractions were as low as 0.05), but that they were unable to discriminate the modal speeds of these distributions. Dr. Watamaniuk completed some additional control experiments at Smith-Kettlewell in which he compared two conditions: either the dots changed their speeds at random every frame, or having once been assigned a speed, chosen at random, they maintained the same speed for the duration of the display. Subjects showed identical precision for these two conditions. He then performed some computer simulations to show that a detector that responds to a single dot's trajectory in either condition would be unable to achieve the precision of the human subject; the single-dot strategy was particularly imprecise for the condition in which each dot maintained its randomly-chosen speed. He concluded that the human subject must pool the signals from many dots in order to estimate the mean speed associated with the global percept. A paper describing this work (Watamaniuk and Duchon) has been accepted by Vision Research.

E. The McKee-Watamaniuk Motion Trajectory Study

A single point moving in apparent motion along a linear trajectory is easily detected when presented against a background of similar points in random apparent motion. Since the motion of this single point (the "signal") can be detected even when the spatial-temporal characteristics of the background "noise" is identical from frame-to-frame to that of the signal, the human motion system must integrate the motion signal for many frames within sensory narrowly runed to particular directions of motion. We measured the detectability of a point moving for 500 msec straight through the center of a ten degree field in one of eight directions, spanning 360 degrees, chosen at random; subjects judged whether the signal point was present or not. Detectability was measured as a function of the increasing density of the noise, an operation that necessarily increased the probability of a mis-match between the signal point and the background points. Surprisingly, subjects could readily detect the signal point ($d' \geq 2.0$) when the probability of a mis-match was as high as 38%, assuming nearest-neighbor matching. Small random perturbations in the straightness of the trajectory ("wobble") had no effect on detectability provided that the directional range of the

perturbations did not exceed a bandwidth of 30 degrees. When the motion of the point was broken into small vectors and displayed in random sequence at positions along the trajectory path, detectability decreased significantly. Thus, the ordered sequence, characteristic of natural motion trajectories, appears to enhance the signal within directionally-tuned mechanisms.

F. The Bravo-Watamaniuk Transparency Studies

If half the dots in a random dot cinematogram move upward at a slow speed, and the other half move upward at a fast speed, two transparent planes are seen, a result that might be predicted from the motion parallax produced by objects at different distances from the observer. Discrimination of small changes in the speed of one set of dots is unaffected by the presence of the other dots. However, when the dots alternate *synchronously* between the two speeds so, at any instant only one speed is present, then only one surface is seen. For all tested alternation rates in this *synchrony* condition, discrimination of either speed is greatly impaired. When the dots alternate *asynchronously* between the two speeds, so at any instant both speeds are present, then two transparent surfaces that "twinkle" are seen. With the change in speed, the dots appear to shift abruptly from one perceptual plane to the other, even though physically each dot is moving along its initial trajectory but at a different speed; the "twinkling" is the perceptual indicator associated with the abrupt disappearance from one of the two planes. In this *asynchronous* condition, discrimination of one speed is unimpaired by the presence of the other speed at all but the fastest alternation rates.

G. Publications

Bowne, S.F., McKee, S.P. & Glaser, D.A. (1989) Motion interference in speed discrimination. Journal of the Optical Society of America, 1112-1121.

McKee, S P. & Welch, L. (1989) Is there a constancy for velocity? Vision Research, 29, 553-561.

McKee, S. P. (1991) Neural coding of local and global motion. Current Biology, 1, 97-98.

McKee, S.P. (1991) The physical constraints on visual hyperacuity. In Vision and Visual Dysfunction, Vol. 5,The Limits of Vision, Kulikowski, J.J., Murray, I.J. and Walsh, V. (Ed.) London: Macmillan.

McKee, S P. & Welch, L (1992) The precision of size constancy. Vision Research, (In Press).

Watamaniuk, S.N.J. & Duchon, A.(1992) The human visual system averages speed information Vision Research, (In Press).

Welch, L. (1989) The perception of moving plaids reveals two motion-processing stages. Nature, 337 734-736.

Welch, L. (1990) Aspects of motion processing revealed by the perception of moving patterns. Doctoral Dissertation submitted in partial fulfillment of the requirements for a Ph.D. from the University of California, Berkeley.

Welch, L. & Bowne, S.F. (1991) Coherence determines speed discrimination. Perception, 425-435.

H. Abstracts

Bowne, S.F. & McKee, S.P. (1988) Contrast discrimination cannot explain speed discrimination. Supplement to Investigative Ophthalmology and Visual Science, 29, 250.

McKee, S.P. & Nakayama, K. (1988) Velocity integration along the trajectory. Supplement to Investigative Ophthalmology and Visual Science, 29, 266

Welch, L. (1988) Speed discrimination and the aperture problem. Supplement to Investigative Ophthalmology and Visual Science, 29 , 264

Welch, L.& Bowne,S.F. (1989) Neural rules for combining signals from moving gratings Supplement to Investigative Ophthalmology and Visual Science, 30, 75.

Welch, L. (1990) Transparency and coherence in moving patterns. (Invited Talk at the Annual Meeting of the Optical Society of America) in Optical Society of America, Technical Digest.

McKee, S.P. & Watamaniuk, S.N.J. (1991) Detecting a single point moving on a linear trajectory amidst randomly moving points. Supplement to Investigative Ophthalmology and Visual Sciences , 892

Bravo, M. J. & Watamaniuk, S. N.J. (1992) Speed segregation and transparency in random dot displays. Submitted to ARVO.

Grzywacz, Norberto M. (1992) One-path model for contrast-independent perception of Fourier and non-Fourier motions. Submitted to ARVO.

Watamaniuk, S. N.J. (1992) Simultaneous direction information from global flow and a local trajectory component. Submitted to ARVO.

Watamaniuk, S. N.J. & Bravo, M.J. (1992) Transparency influences speed discrimination in random dot displays. Submitted to APS.

Motion interference in speed discrimination

Samuel F. Bowne and Suzanne P. McKee

Smith-Kettlewell Eye Research Foundation, 2232 Webster Street, San Francisco, California 94115

Donald A. Glaser

Department of Molecular Biology, University of California, Berkeley, Berkeley, California 94720

Received March 25, 1988; accepted February 7, 1989

Human speed discrimination can be degraded by additional stimuli in close spatial and temporal proximity to the designated test target. In these experiments, observers judged the relative asynchrony between a pair of briefly flashed dots: speed discrimination for two-dot apparent motion. The addition of two irrelevant (interfering) flashed dots to the stimulus, which produces accelerating apparent motion, impaired speed discrimination. We call this impairment motion interference; adjacent stimuli are not processed independently by the motion system. Motion interference is time selective; interfering dots simultaneous with the target dots do not impair speed discrimination, nor do interfering dots that precede or follow the target by 200 msec or more. Motion interference was observed even when the interfering dots were as far away as 1 deg from the test pair. Similar effects were observed with a smoothly moving test target and with interfering stimuli composed only of high spatial frequencies. A multiple-independent-channel model containing several parallel motion-energy detectors with different receptive-field sizes is considered and rejected. We conclude that speed discrimination depends on a time-selective combination of local motion signals from many detectors. These aggregate detectors combine information from local subunits, degrading information about acceleration.

INTRODUCTION

In this paper we address a basic issue for all multiple-channel models of motion detection: independence of channels. Our experiments are essentially masking experiments, in which the speed discrimination for a two-dot target is impaired by adding nearby interfering flashed dots. Before describing the experiments, we begin with a discussion of multiple-channel motion-detection theories.

How does the visual system extract motion information from the pattern of light entering the eye? The Hassenstein–Reichardt model,[1] originally proposed for housefly motion detection, has been extended and applied to human vision by many authors.[2–5] In these models, the image is spatially and temporally filtered, producing two time-dependent contrast signals that are then compared in either a linear[4,5] or a nonlinear[2,3] manner to produce direction selectivity. The motion-energy detector proposed by Adelson and Bergen[2] is representative of this type of model. The other models contain similar spatial and temporal filtering stages and may in some cases be reduced to mathematical equivalence with the Adelson–Bergen model.[2,3] One unifying characteristic of all these models is the use of localized spatial filters: each detector is sensitive to only a limited region of the image and thus computes a measure of local motion. We use the motion-energy detector for the calculations in this paper, but the conclusions are valid for other detectors as well.

The specification of an elementary motion sensor, such as the motion-energy detector, does not constitute a complete model. A large number of localized motion detectors respond to the stimulus, and higher stages of motion processing (the homunculus) must decide which responses to use for the task at hand and which to ignore. This decision, which van Santen and Sperling[3] called a voting rule, is a necessary part of any psychophysical model. In particular, we consider the independent-channel hypothesis: the homunculus has access to the output of every motion-energy detector and is free to choose any one detector's output as its decision variable.

Figure 1 shows a schematic diagram of the parallel-independent-channel model of motion detection. The stimulus is defined by the contrast, which varies in only one spatial dimension x and in time t. This model is one dimensional; the only two directions of motion possible are rightward and leftward. We do not consider extensions of the model to two dimensions in this paper. The parallel motion-energy detectors have different sizes and spatial positions, but they all share a common pair of temporal response functions. The even-symmetric spatial receptive fields of three representative detectors are shown in Fig. 1 (the odd-symmetric receptive fields have been omitted for clarity).

We use the simplest decision stage, following the Klein–Levi model of spatial vision[6]: the winner-take-all homunculus. We assume that so many different motion-energy detectors are present, with a wide variety of positions and sizes, that they form a continuum, densely sampling both space and spatial frequency (SF). The homunculus finds the detector that is most useful for the task at hand and then bases its decision on the output of that single unit, ignoring all others.

Consider a simple two-dot apparent-motion stimulus, in which a dot is flashed briefly and followed 25 msec later by another dot 20 arcmin away. Observers report a sensation of motion between the dots and have an impression of the apparent speed of this motion. We measure the delay between the flashes by stimulus onset asynchrony (SOA). The speed of the apparent motion is given by the (interdot spac-

0740-3232/89/071112-10$02.00

ing)/SOA, or 20 arcmin/25 msec = 13.3 deg/sec. In this paper we are concerned primarily with speed discrimination: the ability of an observer to distinguish a fast motion (short SOA) from a slower motion (longer SOA). Although the speed could also be varied by changing the spatial separation of the dots, we manipulated speed only by means of stimulus timing.

How is two-dot speed discrimination performed by the model shown in Fig. 1? Several motion detectors respond to this stimulus, ranging from small detectors, which see only one dot, to larger detectors, which see both dots. A detector that sees only one dot has no speed information and is therefore useless for speed discrimination. This is a consequence of the winner-take-all homunculus: The homunculus cannot compare the outputs of two small detectors but must attend to only one detector. Note that each motion detector has paired even and odd spatial filters, as shown in Fig. 2. Detectors that see both dots will respond with a burst of motion energy after the stimulus presentation, and the total integrated response will depend on the speed of the apparent motion. The homunculus performs speed discrimination by measuring the output of one of these speed-dependent detectors.

Figure 2 shows the two target dots T and the spatial receptive fields of a motion-energy detector that is sensitive to the apparent speed of the T-to-T apparent motion. For the moment, ignore dots A and B. The details of the calculation showing that this detector is the most sensitive are rather involved and are shown in Appendix A. However, it is easy to understand why this detector is best on a qualitative level. Small detectors are not useful for speed discrimination, because they cannot see both T dots. Large detectors have speed-dependent responses, but the absolute sensitivity of large detectors to small dots declines at low SF's. Burr and Ross[7] showed that contrast sensitivity to moving sine waves is independent of SF over a large range. However, sine-wave stimuli are extended in two dimensions and fill the entire receptive field of each detector. The small dots used in our stimuli occupy a tiny portion of the receptive field of a motion unit, so that the response of a large motion unit falls off as the square of the unit's peak SF. The

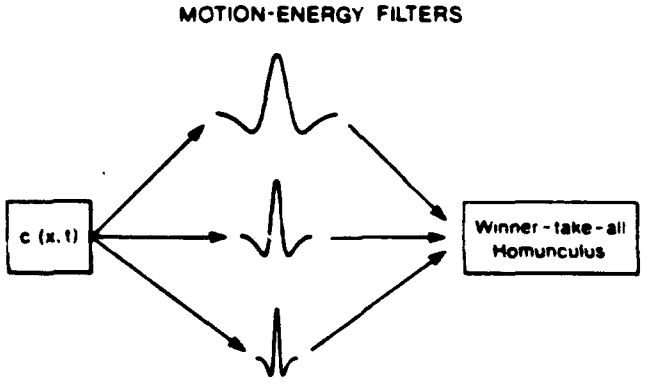


Fig. 1. Schematic diagram of the multiple-independent-channel model of speed discrimination. The input is the contrast, which varies in space and time. This input is processed independently by many motion-energy detectors, of which three are shown. Each motion detector has a response that depends on the stimulus timing. The homunculus chooses the most sensitive motion-energy detector and bases its discrimination entirely on the output of that unit.

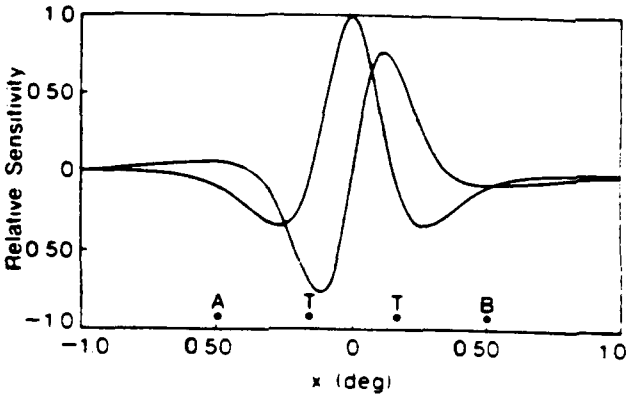


Fig. 2. Spatial receptive fields of a motion-energy detector with σ = 22 arcmin (peak SF 1.3 cpd) are shown, superimposed upon the four-dot motion-interference stimulus. This detector is quite sensitive to the relative timing of the two T dots. Dots A and B are on the edge of this detector's receptive field and have little effect on its response. Both spatial functions contribute equally to the response, as detailed in Appendix A. If speed discrimination were mediated by independent motion-energy detectors of this size, there would be no threshold elevation in the four-dot experiment, regardless of the timing of dots A and B.

detector shown in Fig. 2 is a compromise between these two extremes and has a peak SF response at 1.65 cycles per degree (cpd).

Now consider the effect of adding two more flashed dots, A and B, to the display. As shown in Fig. 2, the optimum detector for speed discrimination is rather insensitive to these dots, which fall on the outskirts of its spatial filters. Therefore dots A and B should not impair speed discrimination, regardless of their timing. The experiments reported below measure the extent to which added dots impair speed discrimination, a phenomenon that we call motion interference.

METHODS

Dot stimuli were computer generated on a cathode-ray-tube (CRT) screen with P4 phosphor. The refresh rate varied but was always greater than 3000 Hz. All times were accurate to ±0.3 msec. Each dot had a presentation time of 1 msec and a diameter of approximately 0.3 mm. The brightness of these dots was measured with a Pritchard UBD 1-deg photometer, by repeating the 1-msec flashes at a repetition rate of 16 Hz and measuring the average luminance of defocused images of the dots. The brightness of each dot was 1 cd, although the dots appeared much dimmer than a continuous 1-cd source because of the 1-msec exposure time. The CRT was illuminated by a tungsten lamp, providing a greenish background illumination of 1.4 cd/m^2. Observers sat at a distance of 91 cm from the CRT screen.

Bar stimuli were shown on a CRT screen with P31 phosphor, using a raster with a refresh rate of 333 Hz. The mean luminance was 16 cd/m^2, and the screen subtended 6 × 5 deg of visual angle at the viewing distance of 144 cm.

The authors Bowne and McKee and four paid university students served as observers. All observers had normal or corrected-to-normal vision, except observer SM, who was an uncorrected hyperope (0.5 D) with 20/20 (Snellen acuity) vision at the viewing distance used in this study. Observers

viewed the stimuli binocularly with natural pupils. The fixation target was presented continuously and was either a dim white dot drawn by the CRT beam or a black mark on the CRT face. The fixation point was always centered vertically on the target and 20 arcmin to the left of the leftmost target dot. The room was dimly lit by fluorescent lights, rendering the frame of the CRT and other objects in the room visible to the subject.

Control experiments showed that reducing the background illumination by more than a factor of 100, reducing target and background luminances simultaneously by a factor of 10, or adding 2 D of optical blur had no effect on speed discrimination with the dot targets.

PSYCHOPHYSICAL PROCEDURE

Speed-discrimination thresholds were determined by the method of single stimuli.[8] We shall describe the procedure for the two-dot stimulus described in the Introduction. Each dot was flashed briefly, with a presentation time of 1 msec. The pair of flashed dots gave rise to a sensation of motion, which the observer classified as fast or slow. No reference stimulus was presented; observers were asked to judge speed relative to the average of the presented speeds. Each trial's asynchrony was chosen randomly from a set of five or seven evenly spaced asynchronies, with upward and downward motions randomly interleaved. After each trial, audible error feedback was given. The first 20 trials were practice trials, intended to define the average speed, and they were followed by 150 trials that were used to obtain the threshold. These 150 responses were accumulated to form a psychometric function: the percentage of fast responses at each asynchrony. This psychometric function was then fitted to a cumulative normal distribution by probit analysis.[9] The asynchrony increment threshold was defined as the increment in target SOA that was required to raise the percentage of fast responses from 50% to 75% on the fitted cumulative normal curve ($d' = 0.67$). Each reported threshold is based on at least 300 responses. The same procedure was used for the four-dot experiments.

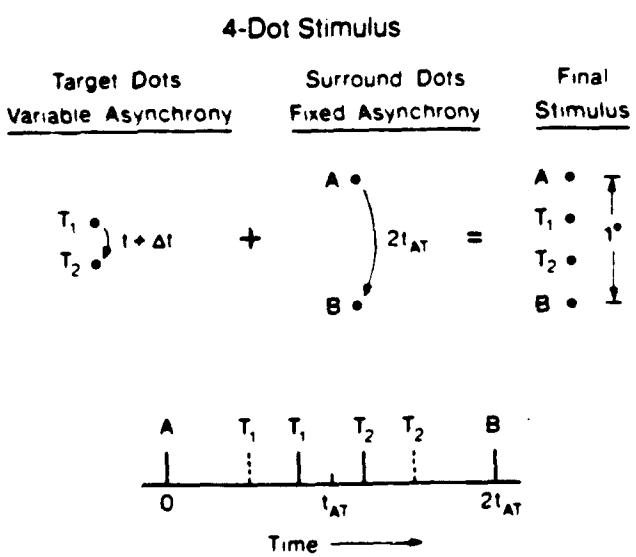


Fig. 3. Four-dot motion-interference stimulus. Two target dots T_1 and T_2 with variable asynchrony $t + \Delta t$ are flanked by two surround dots A and B with a fixed asynchrony $2t_{AT}$, where t_{AT} is the average surround–target asynchrony. Within each block of trials, t_{AT} and t are held constant; only Δt varies. The observer's task is to determine the apparent speed of the T_1-to-T_2 motion, ignoring dots A and B. The time line shows the presentation time of each dot. The solid lines T_1 and T_2 show the onset times when the T_1-to-T_2 asynchrony is $t - \Delta t$, corresponding to fast target motion. The dashed lines show a slow trial. The asynchrony increment threshold is that Δt producing a just-noticeable difference in the apparent target speed.

The primary stimulus was the four-dot stimulus shown in Fig. 3. Four dots are arranged in a vertical line, evenly spaced 20 arcmin apart. Each dot has the same brightness and duration (1 cd and 1 msec). The inner two dots are the target dots (T_1 and T_2) and have an onset asynchrony of $t \pm \Delta t$. The average asynchrony t determines the median target speed, and five or seven values of Δt were used to determine the asynchrony increment threshold. The outer two dots are the surround dots A and B, which have a fixed asynchrony $2t_{AT}$, where t_{AT} is the average surround–target asynchrony. The observer's task is to judge the speed of the T_1-to-T_2 motion, disregarding dots A and B. The values of t and t_{AT} are held fixed during an experimental session, and only Δt is varied, as shown on the time line. Since the timing of dots A and B does not change between the fast and slow displays, dots A and B provide no information useful for this task.

RESULTS

The results of the four-dot experiment are shown in Fig. 4. The T_1-to-T_2 asynchrony increment threshold, Δt, is shown for a range of t_{AT} values. Note that the threshold is elevated when the apparent speed of the A-to-B motion ($1 \text{ deg}/2t_{AT}$) is near the median target speed (13.3 deg/sec). For short or long t_{AT}, the threshold decreases to a value near the threshold obtained from a display containing only dots T_1 and T_2. Evidently, when the surround dots are separated well in time from the target dots, they have no effect on speed discrimination. However, when the target and surround dots are flashed 50 msec before and after the target dots, speed discrimination is greatly impaired.

For some values of t_{AT}, the discrimination task was impossible even with the largest asynchrony increment. These data do not provide a threshold estimate, but only a lower bound: the threshold exceeded 30 msec. This region of greatly elevated thresholds is shown as a solid horizontal line at 30 msec in Fig. 4.

In order to determine whether two-dot asynchrony judgments are similar to speed judgments for continuously moving stimuli, we repeated the experiments of Fig. 4 with a smoothly moving target. Instead of two dots T_1 and T_2, the target consisted of 21 dots spaced 1 arcmin apart, each presented for 0.25 msec sequentially so that the motion appeared continuous. The intensity was adjusted so that the moving target dot and the briefly flashed surround dots A and B appeared equally bright. The motion-interference effect was similar to that shown in Fig. 4 but was weaker. The peak threshold for the smooth-motion stimulus was 10–20 msec in the presence of the interfering dots, 2 to 4 times the threshold for the target alone. This qualitative agreement between smooth motion and sampled target motion is consistent with the spatial filtering performed by motion-energy detectors and the equivalence of speed discrimina-

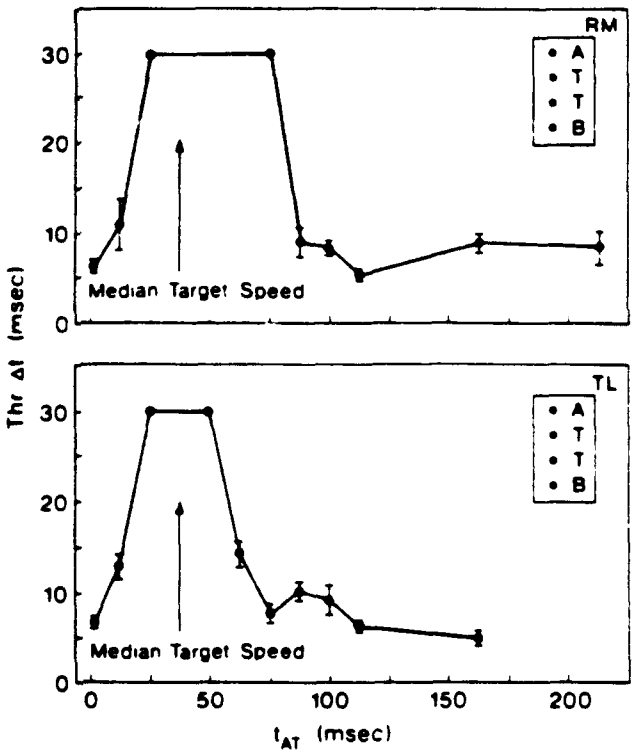


Fig. 4. Results of the four-dot motion-interference experiment for two observers. The asynchrony increment threshold Δt is plotted as a function of the surround–target asynchrony t_{AT}. The average asynchrony between the targets T_1 and T_2 was 25 msec, resulting in a median target speed of 13.3 deg/sec. The arrow indicates the point at which the surround speed (1 deg/2t_{AT}) equals the median target speed. Near the arrow, the thresholds increased to values >30 msec. As explained in the text, only a lower limit on these thresholds was measured.

tion for smooth and sampled stimuli found by McKee and Welch.[8]

The motion-interference effect was also measured with the average asynchrony of the target dots increased to 50 msec. The results were virtually identical to those found with a 25-msec asynchrony, with a maximum threshold elevation near t_{AT} = 40 msec.

Since many previous studies were concerned with direction-of-motion discrimination rather than with speed discrimination, we wondered whether motion interference also affects this judgment. In the next experiment the four-dot stimulus shown in Fig. 3 was used. The SOA of dots A and B was held constant within each session, as before, but the direction of the A-to-B motion was not varied from trial to trial. The T_1-to-T_2 motion had random direction and asynchrony. The observer's task was to judge the direction of motion of the target dots (up or down), ignoring dots A and B. In the absence of the surround dots, this task is easy, and asynchrony thresholds as small as 3–4 msec were reported.[9] The added dots A and B make the direction much more difficult to distinguish, and a large amount of practice was needed (approximately 1000 trials) to obtain the thresholds shown in Fig. 5. Direction-of-motion discrimination is impaired when the surround–target asynchrony t_{AT} is 50 msec or less, with the maximum threshold elevation occurring when the surround and target dots are simultaneous. Similar results were reported by MacLeod *et al.*[10] and Green.[11]

Spatial Range of Motion Interference

Figure 6(a) shows the motion-interference effect with the surround dots 1 deg away from the target dots. The threshold elevation is much smaller and the optimum t_{AT} for interference is much longer than those observed with closer surround dots, but motion interference is still present.

Figures 6(c) and 6(d) show the results of an experiment in which the A-to-B motion is not collinear with the target motion. Even when the vertical motion paths are separated horizontally by 1 deg large interference effects are seen, similar to the effects shown in Fig. 6(a).

Perhaps the most striking example of off-axis motion interference is shown in Fig. 6(b). The target contained 21 closely spaced dots and appeared to be a single smoothly moving dot, and the interfering dot A was flashed for 1 msec. The ordinate t_{AT} is the time between the onset of dot A and the onset of the center target dot. Data were collected with A preceding T and with T preceding A. The thresholds were similar in both cases, and so they were averaged together.

Spatial-Frequency Selectivity

We repeated the four-dot experiment, using bar targets, in order to measure the relative contribution of low- and high-SF motion detectors to motion interference. The four dots shown in Fig. 3 were replaced by vertical bars 5 arcmin wide and 5 deg long, as shown in Fig. 7(a). The timing was the same as that shown in Fig. 3; observers were asked to attend to the taret bars and to ignore the interfering bars A and B. The contrast of the target bars T was 2.5 times threshold for observer SFB and 3 times threshold for observer SPM. The contrast of the interfering bars A and B was set to 2 times threshold for observer SFB and 1.25 times threshold for observer SPM. Thresholds were determined by a two-temporal-interval forced-choice QUEST staircase[12] and defined as the 92% correct point on the psychometric function (d' = 2.0). To reduce the contribution of low-SF motion detectors, we replaced the interfering bars A and B with three-bar high-SF patterns as shown in Fig. 7(b). Features A and B each contained a central bright bar flanked by two dark bars, with the bright/dark ratio adjusted psychophysically to give the minimum detectability as measured by multiple interleaved QUEST staircases. The contrasts of the high-SF A

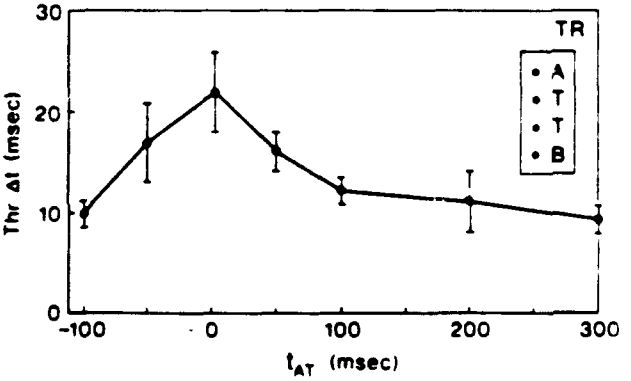


Fig. 5. Direction of motion discrimination with the four-dot stimulus. The order of presentation of the surround dots was AB for t_{AT} > 0 and BA for t_{AT} < 0. The observer's task was to determine the direction of motion of the target dots, while ignoring dots A and B. The threshold is elevated when the surround–target delay is 50 msec or less.

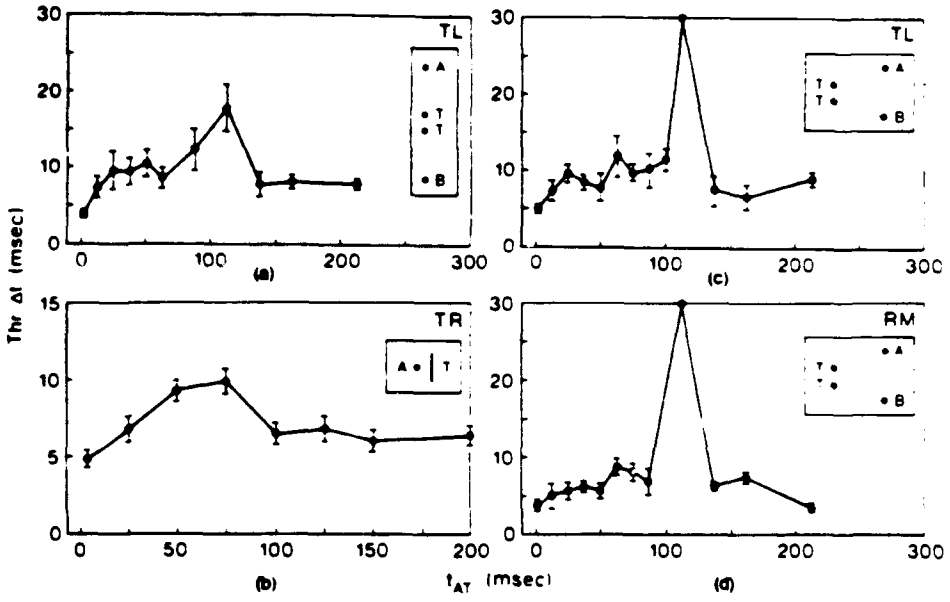


Fig. 6. Motion interference with distant surround dots. (a) Surround dots 1 deg above and below the target dots. (b) Interference by a single surround dot 10 arcmin to the left of a smooth-motion target 20 arcmin long. (c) and (d) Surround dots 20 arcmin above and below and 1 deg to the right of the target dots. (Results are shown for two observers.) The median target speed was 13.3 deg/sec.

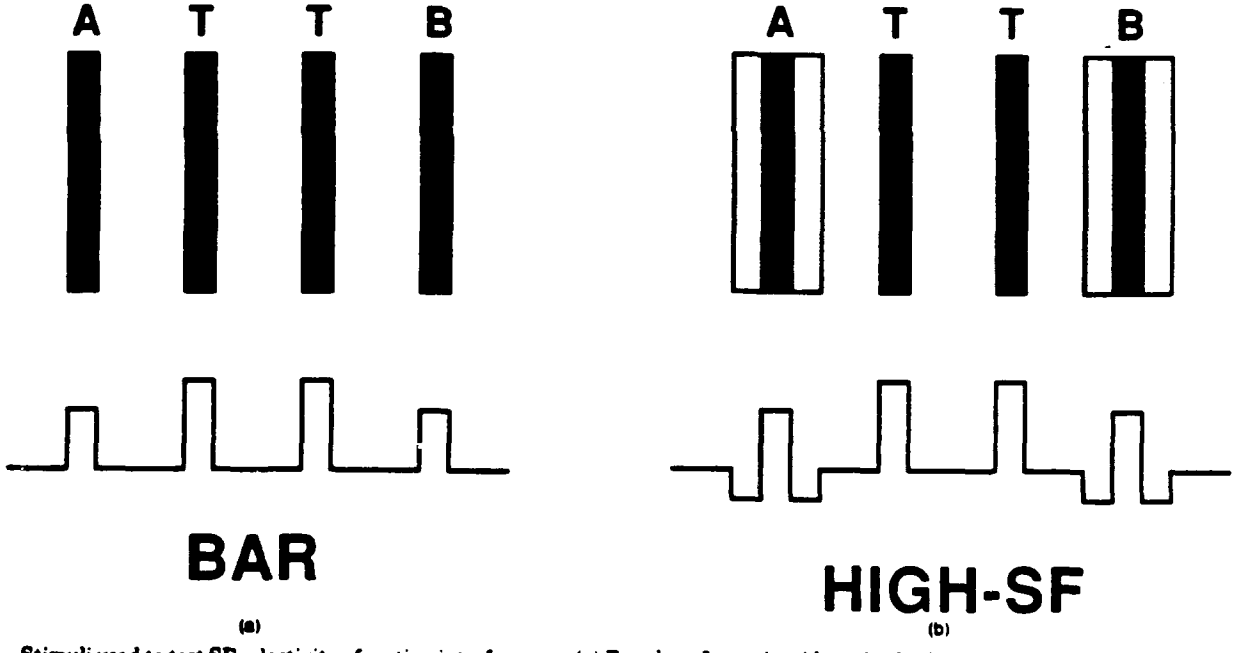


Fig. 7. Stimuli used to test SF selectivity of motion interference. (a) Four bars 5 arcmin wide and 5 deg high are flashed sequentially with the timing shown in Fig. 3. The observer's task is to determine the speed of the T-to-T motion, ignoring bars A and B. The target bars T had contrasts of 2.5 times threshold for observer SFB and 3 times threshold for observer SPM. The interfering bars A and B had contrasts of 2 times threshold for observer SFB and 1.25 times threshold for observer SPM. (b) The interfering targets A and B have been replaced by high-SF patterns containing a central bright bar flanked by two dark bars. The bright/dark ratio was adjusted for minimum visibility, and the contrasts presented were 2 times threshold for observer SFB and 1.25 times threshold for observer SPM. The target contrasts and timing were the same as in (a).

and B patterns were exactly comparable with the bar A and B patterns: 2 times threshold for observer SFB and 1.25 times threshold for observer SPM.

As shown in Fig. 8, the motion interference extends over a larger time range than that shown in Fig. 3 but is qualitatively similar. The main result is that motion interference is substantial in both subjects even with high-SF interfering patterns. The motion interference is equally strong for bar and high-SF patterns for observer SPM and only partially diminished for observer SFB. This result has important implications for models of motion interference, as we discuss below.

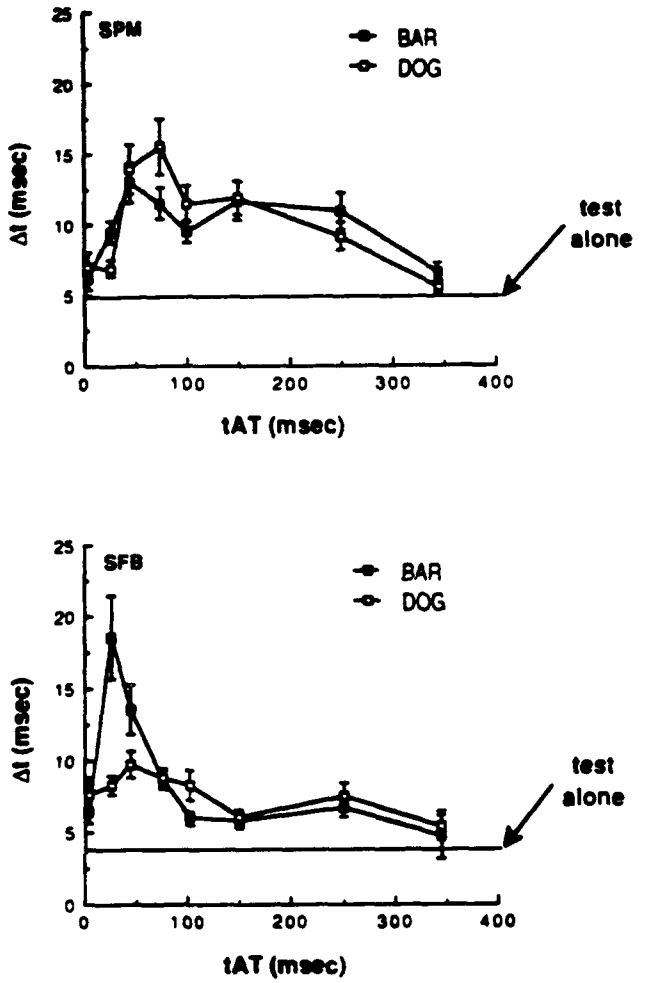


Fig. 8. SF selectivity of motion interference for two observers. Time increment thresholds Δt are shown as a function of T_{AT}, the mean surround–target asynchrony. Filled symbols show thresholds obtained with bars as interfering features, and open symbols show thresholds obtained with high-SF interfering features [difference-of-Gaussian (DOG) functions]. Solid horizontal lines show the threshold in the absence of features A and B. For both observers, thresholds are elevated substantially by both bar and high-SF interfering features. If motion interference were caused by a coarse-to-fine interaction among motion units, the high-SF pattern would cause much less threshold elevation than the bars.

DISCUSSION

The four-dot experiment clearly refutes the model shown in Fig. 1; dots A and B impair speed discrimination even when they are so far away from the target (1 deg) that they should have no effect at all on the motion-energy unit that is most useful for speed discrimination. We now consider other models that might explain motion interference.

Spatial Filter Size

Could speed discrimination in the four-dot stimulus shown in Figs. 2 and 3 be mediated by motion-energy units with very large spatial filters (0.5 cpd or lower) that are unable to resolve the target dots from the interfering dots A and B? This proposal seems unlikely on both theoretical and experimental grounds.

First we present the theoretical argument. Figure 9 shows the speed-discrimination thresholds predicted from the motion-energy model described in Appendix A for several detector sizes. The four-dot stimulus shown in Figs. 2 and 3 was used in these calculations. The horizontal axis is the average time between dot A and the target dots, as defined in Fig. 3, and the vertical axis is the predicted asynchrony increment threshold Δt, in arbitrary units. The predictions for four detector sizes are shown. The two largest detectors, with space constants σ of 1 and 3 deg (peak SF's of 0.48 and 0.16 cpd) are unable to resolve the target dots from dots A and B and therefore show large motion-interference effects. However, the speed-discrimination performance of these large units is poor: the smaller detector with $\sigma = 0.3$ deg (peak SF 1.6 cpd) is much more sensitive to target speed at all t_{AT} values. This smaller unit shows no motion interference at all. It seems unlikely that the outputs of these smaller units are ignored by the homunculus in preference to the less-sensitive larger units.

Could it be that the visual system does not contain motion detectors with spatial frequencies near 1.6 cpd? Such motion detectors have been found by Anderson and Burr, using both masking[13] and subthreshold summation[14] techniques with sine-wave stimuli. Furthermore, McKee and Taylor[15] showed that the optimum spacing of two lines for speed discrimination is 5–10 arcmin in the fovea, which implies that the motion units mediating two-line speed discrimination are not much larger than the unit shown in Fig. 2.

The results of the experiment with the high-SF target show conclusively that motion interference can be detected even when low SF's are excluded from the interfering targets. These results also rule out models that use a coarse-to-

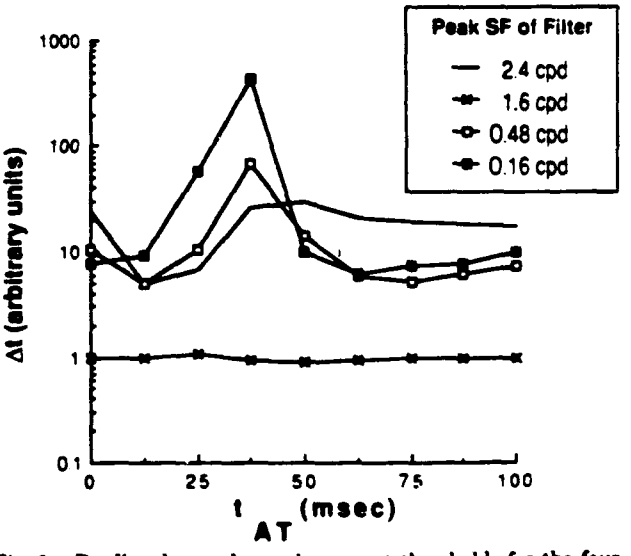


Fig. 9. Predicted asynchrony increment thresholds for the four-dot experiment, for detectors of various sizes. Although the threshold scale is arbitrary, the vertical separations of the curves are accurate. The two lowest-SF detectors (0.16 and 0.48 cpd) show motion-interference effects crudely similar to those shown in Fig. 4, because they are so large that they cannot resolve the target and surround dots. The 1.6-cpd detector is optimal for speed discrimination, having a time increment threshold lower than any other SF. This detector is insensitive to the interfering dots A and B, as shown in Fig. 2. The 2.4-cpd detector is so small that the two target features fall on insensitive regions of its receptive field, so it is not sensitive to target speed.

fine interaction between scales, as proposed in some models of spatial vision,[16,17] since the interference effect is still present in the absence of low-SF signals.

Subunit Organization

We propose that motion units are in fact similar to motion-energy detectors and that speed discrimination is mediated by units near 1.65 cpd for the stimulus shown in Fig. 2 but that motion units are not independent. Instead, individual motion-energy units act as subunits that are combined together to produce a final motion sensor that is stimulated best by a small target moving with constant speed. This combination of motion-energy signals reduces the sensitivity of the system to accelerating motions such as that shown in Fig. 3 but may have other advantages. For example, such a network may be better at detecting a moving object in a noisy environment or it may eliminate local motion signals arising from occlusion or other nonmotion signals.

CONCLUSION

The multiple-independent-channel model cannot explain motion interference. Theoretical and experimental results argue that the two-dot target is detected by a motion unit with a peak SF near 1.65 cpd, but such units are not affected by the interfering dots A and B. Therefore the multiple-independent-channel model cannot explain speed discrimination. This is a surprising result, especially when we consider the success of multiple-independent-channel models in spatial vision.[6,18] In addition, models that mix small-scale information with large-scale information[16-19] cannot explain motion interference, because removing low SF's from the interfering features does not eliminate the effect. Turano and Pantle[20] proposed a two-stage motion model, based on an elegant series of experiments using amplitude-modulated sine-wave gratings, that is qualitatively supported by the motion-interference results.

Results of previous psychophysical studies[10,11] also show that motion detectors interact in a manner that can destroy information. Results of physiological studies[21,22] support this idea. The time dependence of motion interference suggests that the interaction involves combining the outputs of motion detectors from different spatial locations at different times, analogous to a Hassenstein–Reichardt detector whose inputs are motion-energy responses. Similar models were proposed to explain both physiological[23] and psychophysical[24] data. Burr[25] showed that the detectability of a moving dot is greater than the detectability of a flash of light containing the same amount of light distributed over the same region in space and time, which may also be a consequence of the nonindependence of motion units.

The motion-interference results suggest that the time delay for the proposed second stage of comparison is approximately 50 msec. The long-range results shown in Fig. 6 suggest that the effective delay is longer when the interfering dots are farther from the target dots, which may be a consequence of recurrent delayed inhibition in the second-stage processing.

Finally, we should discuss the purpose of this subunit structure. Noise, occlusions, transparency, specular reflection, and failure to solve the correspondence problem all may contribute spurious motion signals in local detectors, which will diminish the visibility of moving objects. One way to remove these unwanted signals is to enforce consistency across spatial scale,[16,17] a scheme particularly suited for static spatial patterns. However, when objects move, a new regularity in the image emerges if the moving objects have constant speed. Combining information along the trajectory of a moving object provides an independent check on the data provided by local motion detectors, which may be used instead of, or in addition to, the comparison of images across scale.

APPENDIX A: MOTION-ENERGY VIEWPRINT CALCULATIONS

Our motion model is an extension of the viewprint model of Klein and Levi,[6] which they used to explain the results of hyperacuity experiments. We use the motion-energy detectors of Adelson and Bergen.[2] Our model assumes that a large number of motion-energy detectors with different sizes and positions respond independently to the stimulus, as shown in Fig. 1. Each detector has a speed-dependent response; in principle, a number of detectors could contribute to speed discrimination. However, we assume that discrimination is done with a single motion-energy unit, just as in the viewprint model.[6] Since the discrimination tasks were done with feedback and after practice, we assume that the observer had learned to use that detector that is most sensitive to the asynchrony of the target dots.

First, we summarize the calculation of the response of a single channel: one motion-energy detector. The procedure is exactly that of Adelson and Bergen, except that Cauchy space functions are used instead of Gabor or Gaussian-derivative functions. The model is one dimensional, so only two directions of motion are possible: leftward and rightward.

As shown in Fig. 2, two overlapping receptive fields are used: a symmetric function $S(x)$ and an antisymmetric function $A(x)$. These functions represent the relative sensitivities of the detector to light at different spatial positions. The particular spatial functions that we use (Cauchy 3 functions[6]) are shown in Fig. 2. The Cauchy functions were chosen because of mathematical convenience; the results would be similar if difference-of-Gaussian functions or Gabor functions were used. A detailed discussion of the properties of various spatial receptive-field functions was given by Klein and Levi[6]; none of the conclusions in this paper depends greatly on this choice. Here are the equations defining the Cauchy functions:

$$S(x) = (1 - 6s^2 + s^4)/(1 + s^2)^4, \tag{A1}$$

$$A(x) = -4s(1 - s^2)/(1 + s^2)^4, \tag{A2}$$

where $s = x/\sigma$. σ is a measure of receptive-field size or SF tuning. These functions are bandpass in SF, peaking at a SF of $3/(2\pi\sigma)$. For example, when $\sigma = 0.33$ deg = 20 arcmin, the peak SF is 1.4 cpd.

Two overlapping temporal decay functions are used, $f_3(t)$ and $f_5(t)$:

$$f_n(t) = (kt)^n \exp(-kt)[1/n! - (kt)^2/(n + 2)!], \tag{A3}$$

where $n = 3$ or $n = 5$.

Two separable directionally selective filters r_1 and r_2 are defined as follows:

$$r_1(x, t) = S(x)f_5(t) - A(x)f_3(t),$$
$$r_2(x, t) = A(x)f_5(t) + S(x)f_3(t), \quad \text{(A4)}$$

where x is the location of the filter and t is the time at which the filter response is to be calculated. Contour plots of r_1 and r_2 are shown in Fig. 10.

These are linear filters; the response of the filter centered at x_c at time t is given by

$$R_1(x_c, t) = \int r_1(x - x_c, t - t')c(x, t')\mathrm{d}x\mathrm{d}t', \quad \text{(A5)}$$

where $c(x, t)$ is the stimulus contrast.[26]

These filters respond more to rightward motion than to leftward motion, but, since they are linear, they are phase sensitive: the response depends on the sign of the stimulus contrast, and a drifting sine-wave grating elicits a sinusoidally oscillating response. To remove phase sensitivity, we square and sum these responses. We then take the square root to produce time-dependent rightward-motion energy, $R(x, t)$:

$$R(x, t) = [R_1(x, t)^2 + R_2(x, t)^2]^{1/2}. \quad \text{(A6)}$$

Unlike R_1 and R_2, R is always positive and reaches a constant value when the stimulus is a sine-wave grating drifting at constant speed. It also responds similarly to a moving dark bar and to a moving light bar. R is proportional to the stimulus contrast.

The stimuli that we used were all brief presentations, so that $R(x, t)$ increases from zero to a peak value and then decreases to zero after the stimulus presentation ends. We assume that speed discrimination is determined not by this entire function but by the total time-integrated rightward motion energy, following the formalism of van Santen and Sperling.[27] The time-integrated motion energy $R(x)$ is calculated as follows:

$$R(x) = \int \mathrm{d}t R(x, t). \quad \text{(A7)}$$

This time-integration assumption is justified better for the brief presentations, when the entire display is within the critical duration of Bloch's law[28] (35 msec), than for the longer presentations. However, we use this integral for all the stimuli modeled, for simplicity. Watson[29] presented a model for probability summation over time in which the time integral is calculated over all time, but those regions of time in which the response is large are weighted more highly. However, this model was developed to explain contrast sensitivity, when the target is at the detection threshold. For speed discrimination of suprathreshold targets, the observer may be able to attend to some restricted period of time, ignoring responses at other times. It is therefore not clear whether the weighting used by Watson is appropriate here. We therefore used the simple time-integral formulation, assuming that the observer attends equally to all instants of time after our brief presentation. The use of other temporal weighting schemes such as Watson's would not change the conclusions.

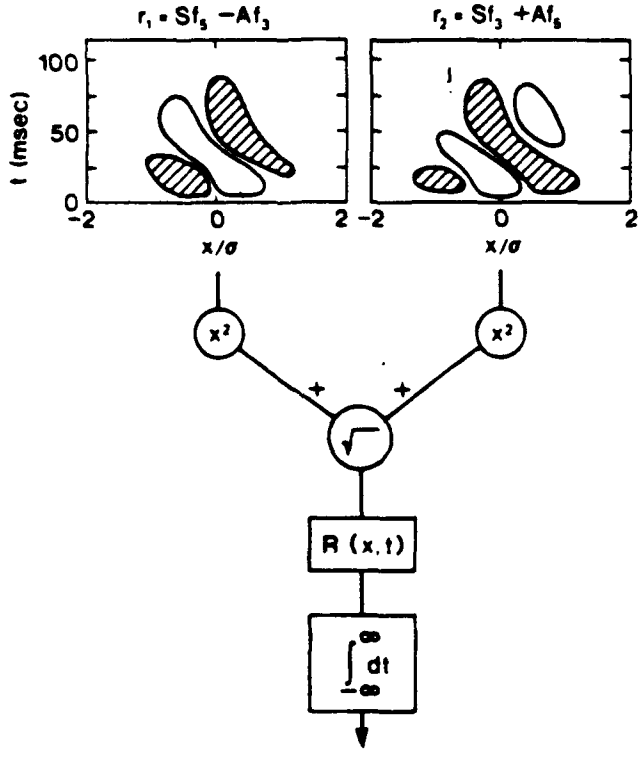


Fig. 10. Motion-energy calculations. At the top, space–time contour plots of the filters in the model are shown. In the hatched regions, the responses are negative, and in the unshaded regions, the responses are positive. r_1 and r_2 have some direction selectivity but are phase sensitive. $R(x, t)$ measures the rightward motion energy at any instant t, for the detector centered at x. $R(x, t)$ is time integrated over the whole epoch of its response to a stimulus, and the resulting $R(x)$ is used for speed discrimination.

Relative Sensitivities of the Motion Detectors

We now need to determine the relative sensitivity of motion detectors tuned to different SF's. We measured contrast sensitivity for vertical sine waves drifting horizontally at a temporal frequency of 8 Hz, windowed by a temporal Gaussian with a standard deviation of 250 msec. The mean screen luminance was 18 cd/m^2, and patterns were refreshed at 100 Hz by a Picasso image generator. The pattern was 3 cycles wide and 2.5 cycles high at every SF tested. Thresholds were determined by using a QUEST staircase[13] with two temporal intervals and were measured at the 92% correct level ($d' = 2.0$). These data agree well with previous measurements by Burr and Ross.[7] Figure 11 shows the measured contrast sensitivities for two subjects and the SF profiles of several representative motion-energy units.

The rate constant k was determined by comparison with the data of Burr and Ross[7] to be 150/sec and is in rough agreement with the temporal impulse response as measured by two-pulse summation.[30,31]

The contrast sensitivity of a single motion-energy detector was calculated by assuming that the difference between the squared rightward and leftward motion energies must reach a threshold value for identification of the direction of motion, in order to mimic the psychophysical criterion used by Burr and Ross. The difference of squared energies was chosen for mathematical convenience and because it approximates the accelerating nonlinearity underlying con-

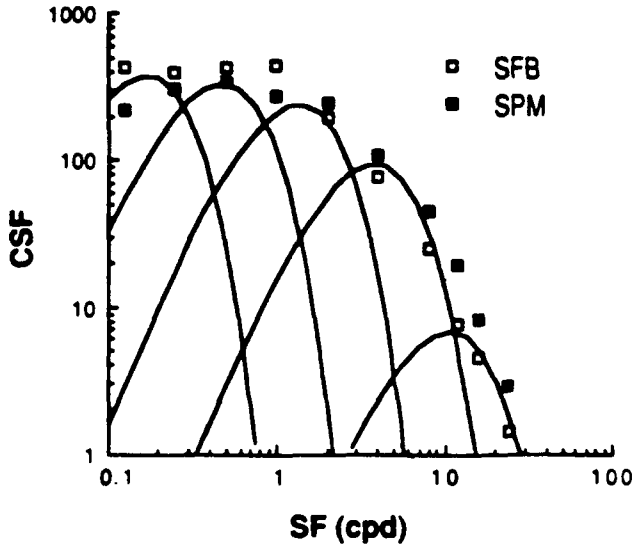


Fig. 11. Contrast sensitivities (CSF's) of two observers to an 8-Hz drifting sine-wave target are shown along with the sensitivity profiles of several motion-energy detectors. The relative sensitivities of the detectors have been adjusted to fit the data and are used in the calculation of the results in Fig. 9.

trast detection.[32] The steady-state contrast-sensitivity function has the following form:

$$\mathrm{CSF}(f_s, f_t) = G(2\pi f_s\sigma)^3 \exp(-2\pi f_s\sigma)P(\omega),$$

$$P(\omega) = \omega^{3/2}(\omega^{14} + 10\omega^{12} + 39\omega^{10} + 80\omega^8 + 95\omega^6 + 66\omega^4 + 25\omega^2 + 4)^{1/2}(1 + \omega^2)^{-7},$$

$$\omega = 2\pi f_t/k. \tag{A8}$$

The sine-wave stimulus has a SF of f_s cpd and a temporal frequency of f_t cycles/sec. The motion-energy detector has a Cauchy spatial parameter σ in degrees and a rate parameter k in reciprocal seconds. G is an arbitrary constant. We assume that many motion-energy filters with equal peak sine-wave sensitivities respond in parallel to the stimulus, so the observed contrast-sensitivity function will be proportional to $P(\omega)$ and will depend only on temporal frequency. The solid lines in Fig. 11 are plots of the function $P(\omega)$, with the maximum sensitivity adjusted to fit the data.

Speed Discrimination with a Single Detector

As an illustration, we describe the speed-discrimination calculation for the four-dot stimulus shown in Fig. 3. The integrated motion energy $R(x)$ depends on the target speed and may be used as a measure of that speed. We calculate $R_-(x)$, the energy for a presentation with the target asynchrony equal to $t - \Delta t/2$, where t is the mean asynchrony. We then calculate $R_+(x)$, the energy for a presentation with target asynchrony $t + \Delta t/2$, where Δt is the asynchrony increment. The energy increment is $dR(\Delta t) = R_+(x) - R_-(x)$. If the energy increment $dR(\Delta t)$ is large, the asynchrony increment is detected by the motion-energy detector. If $dR(\Delta t)$ is small, the asynchrony increment is not detected. For these calculations, we used a fixed value of dR_{thr}, the threshold increment in R. At present we have no experimental estimate of dR_{thr}, so it remains as an unspecified parameter. The uncertainty of dR_{thr} is the only reason that the y axis in Fig. 9 is in arbitrary units.

As described above, we calculate $dR(\Delta t)$, the energy increment produced by an asynchrony increment Δt. We define the asynchrony increment threshold as $\mathrm{Thr} = \Delta t\, dR_{thr}/dR(\Delta t)$. If we assume that the $dR(\Delta t)$ function is locally linear, then Thr is the asynchrony increment that produces a threshold change in R.

The asynchrony increment thresholds in Fig. 9 were calculated in this manner, with various values of σ. The same value of dR_{thr} was used throughout, so the relative sensitivities of the various detectors were preserved.

We considered determining dR_{thr} from contrast increment thresholds, such as those determined by Legge and Foley,[33] assuming that the same mechanism determines both contrast discrimination and speed discrimination. However, we have measured contrast increment thresholds for drifting gratings and found that they do not explain speed-discrimination thresholds.[34] Whereas contrast discrimination thresholds are a power function of background contrast, speed discrimination is independent of contrast at high contrasts.[35] Therefore having a single fixed dR_{thr} is both simpler and more in accord with experimental results than setting dR_{thr} from contrast-discrimination data would be.

ACKNOWLEDGMENTS

We thank Brent Beutter, T. Kumar, and Leslie Welch for many helpful discussions about motion perception and psychophysical experiments. This research was supported by National Institutes of Health grant F32 EY06506 to S. F. Bowne, U.S. Air Force Office of Scientific Research grant AFOSR-85-0380 to S. P. McKee, and U.S. Office of Naval Research grant ONR N000 14-85-K-0692 to D. A. Glaser. S. F. Bowne was also supported by a Miller Institute Fellowship.

REFERENCES AND NOTES

1. W. Reichardt, "Autocorrelation, a principle for the evaluation of sensory information by the central nervous system," in *Sensory Communication*, W. A. Rosenblith, ed. (MIT Press, Cambridge, Mass., 1961); B. Hassenstein and W. Reichardt, "Systemstheoretische Analyse der ZeitReihenfolgen- und Vorzeichanauswertung bei der Bewegegungsperzeption des Russelkafers *Chlorophanus*," Z. Naturforsch. B 11, 513–524 (1956).
2. E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," J. Opt. Soc. Am. A 2, 284–299 (1985).
3. J. P. H. van Santen and G. Sperling, "Elaborated Reichardt detectors," J. Opt. Soc. Am. A 2, 300–321 (1985).
4. A. B. Watson and A. J. Ahumada, Jr., "Model of human visual-motion sensing," J. Opt. Soc. Am. A 2, 322–342 (1985).
5. A. Pantle and K. Hicks, "Using low-level filters to encode spatial displacements of visual stimuli," Spatial Vision 1, 69–82 (1985).
6. S. A. Klein and D. M. Levi, "Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation," J. Opt. Soc. Am. A 2, 1170–1190 (1985).
7. D. C. Burr and J. Ross, "Contrast sensitivity at high velocities," Vision Res. 22, 479–484 (1982).
8. S. P. McKee and L. Welch, "Sequential recruitment in the discrimination of velocity," J. Opt. Soc. Am. A 2, 243–251 (1985).
9. D. J. Finney, *Probit Analysis* (Cambridge U. Press, Cambridge, 1971).
10. D. A. Macleod, L. Welch, and S. P. McKee, "Local interactions affecting direction and velocity of apparent motion," Invest. Opthalmol. Vis. Sci. Suppl. 26, 189 (1985).

11. M. Green, "Visual masking by flickering surrounds," Vision Res. 23, 735–744 (1983).
12. A. B. Watson and D. G. Pelli, "QUEST: A Bayesian adaptive psychometric method," Percept. Psychophys. 33, 113–120 (1983).
13. S. J. Anderson and D. C. Burr, "Spatial and temporal selectivity of the human motion detection system," Vision Res. 25, 1147–1154 (1985).
14. S. J. Anderson and D. C. Burr, "Receptive field size of human motion detection units," Vision Res. 27, 621–635 (1987).
15. S. P. McKee and D. G. Taylor, "Discrimination of time: comparison of foveal and peripheral sensitivity," J. Opt. Soc. Am. A 1, 620–627 (1984).
16. D. Marr and E. Hildreth, "Theory of edge detection," Proc. R. Soc. London Ser. B 207, 187–217 (1980).
17. R. J. Watt and M. J. Morgan, "A theory of the primitive spatial code in human vision," Vision Res. 25, 1661–1674 (1985).
18. H. R. Wilson, "Responses of spatial mechanisms can explain hyperacuity," Vision Res. 26, 453–469 (1986).
19. G. B. Henning, B. G. Hertz, and D. E. Broadbent, "Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency," Vision Res. 15, 887–897 (1975).
20. K. Turano and A. Pantle, "On the mechanism that encodes the movement of contrast variations: velocity discrimination," Vision Res. 29, 207–222 (1989).
21. J. Allman, F. Miezen, and E. McGuiness, "Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT)," Perception 14, 105–126 (1985).
22. B. J. Frost and K. Nakayama, "Single visual neurons code opposing motion independent of direction," Science 220, 744–745 (1983).
23. H. B. Barlow and W. R. Levick, "The mechanism of directionally selective units in the rabbit's retina," J. Physiol. 178, 477–504 (1965).
24. A. M. M. Lelkins and J. J. Koendernik, "Illusory motion in visual displays," Vision Res. 24, 1083–1090 (1984).
25. D. C. Burr, "Temporal summation of moving images by the human visual system," Proc. R. Soc. London Ser. B 211, 321–339 (1981).
26. The stimuli in this paper were all small, bright points or lines. The contrast function is then a sum of Dirac delta functions, and this integral reduces to a finite sum. The quantitative value of the peak contrast requires careful definition for such stimuli, since the bright dots may change the effective background luminance. However, for our purposes the peak contrast is irrelevant, since we calculate only relative thresholds. The calculation of absolute speed-discrimination thresholds is beyond the scope of this paper. Klein and Levi[6] resolved this issue by defining a "local contrast" measure, which could be extended to include temporal contrast variations.
27. J. P. H. van Santen and G. Sperling, "Temporal covariance model of human motion perception," J. Opt. Soc. Am. A 1, 451–473 (1984).
28. A. Gorea and C. W. Tyler, "New look at Bloch's law for contrast," J. Opt. Soc. Am. A 3, 52–61 (1986).
29. A. B. Watson, "Probability summation over time," Vision Res. 19, 515–522 (1979).
30. A. B. Watson and J. Nachmias, "Patterns of temporal interaction in the detection of gratings," Vision Res. 17, 893–902 (1977).
31. J. R. Bergen and H. R. Wilson, "Prediction of flicker sensitivities from temporal three-pulse data," Vision Res. 25, 577–582 (1985).
32. J. M. Foley and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision," Vision Res. 21, 1041–1053 (1981).
33. G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. 70, 1458–1471 (1980).
34. S. F. Bowne and S. P. McKee, "Contrast discrimination cannot explain speed discrimination," submitted to Invest. Ophthalmol. Vis. Sci. Suppl.
35. S. P. McKee, G. H. Silverman, and K. Nakayama, "Precise velocity discrimination despite random variations in temporal frequency and contrast," Vision Res. 26, 609–619 (1986).

Vision Res. Vol. 29, No. 5, pp. 553–561, 1989
Printed in Great Britain.

0042-6989/89 $3.00 + 0.00

IS THERE A CONSTANCY FOR VELOCITY?

SUZANNE P. MCKEE and LESLIE WELCH
Smith–Kettlewell Eye Research Institute, 2232 Webster St, San Francisco, CA 94115, U.S.A.

(*Received* 22 *March* 1988; *in revised form* 12 *September* 1988)

Abstract—Human observers are unable to use disparity information to transform the angular velocity signal into a precise object-based code. The Weber fraction for discriminating changes in objective velocity (cm/sec) is about twice the Weber fraction for discriminating changes in angular velocity (deg/sec), and is substantially higher than predicted from a combination of the errors in judging disparity and angular velocity. By comparison, judgments of the distance traversed by the moving target show excellent size constancy. The discrimination of changes in objective size (cm) is as precise as the discrimination of changes in angular size (deg). The angular velocity signal is useful without transformation into an object-centered signal; it guides eye and body movements, and is the basis of motion parallax judgments. The need to retain this angular signal may explain why there is no efficient mechanism for velocity constancy.

Motion Velocity discrimination Size discrimination Size constancy Velocity constancy

INTRODUCTION

The retinal velocity produced by an object moving at a given physical velocity depends on its distance from the eye. Human observers are generally more interested in the physical dimensions of objects than in the retinal stimuli associated with these dimensions, so it is widely assumed that some compensatory mechanism automatically adjusts angular dimensions by a factor related to perceived distance (Helmholtz, 1868; Woodworth, 1938; Epstein, 1973). If an observer had perfect knowledge about the distance between his head and a moving object, then presumably its perceived velocity would correspond perfectly to the actual physical velocity—constancy would prevail. In this study, we will examine the influence of this presumed compensatory mechanism on velocity *discrimination*.

Typically, constancy studies use matching or magnitude estimation to assess what the observer perceives under various experimental manipulations. In a matching study, there is no way for an observer to be wrong; he is the ultimate arbiter of what he perceives. But in a discrimination study, the experimenter defines what is correct based on the physical characteristics of the stimulus. For example, an observer can be asked to choose the larger of two objects independent of their relative distances, and the experimenter can score the judgments on the basis of the physical size of the objects. If the observer is told about the correctness of his choices ("given feedback"), then the precision of his judgments reveals how well size constancy operates for the tested range of distances. Discrimination judgments do not really tell us what the observer perceives, because an observer may perceive one thing, but respond with another in order to satisfy the experimenter's definition of "correct". Nevertheless, the relative precision of angular and objective judgments can supply some information about how sensory signals are combined to estimate the properties of objects.

Consider two traditional models of constancy. In the first, the "Helmholtz" model of Fig. 1, the observer has access to two independent signals: the retinal signal coded in angular units, and a depth signal. In any plausible biological system, both of these signals are subject to error—they are noisy. The observer achieves constancy by correcting the perceived retinal signal by his estimate of perceived distance, so there are two sources of error in his estimate of the physical dimensions of objects, the retinal error and the depth error. It hardly matters whether the correction is a conscious or "unconscious inference"; the discrimination of objective dimensions is necessarily less precise than the discrimination of angular dimensions. The important assumption of this model is that the observer has access to a signal coded in angular co-ordinates.

In the "Gestalt" model of constancy, the observer has access only to constancy-corrected signals—signals which are already adjusted by a

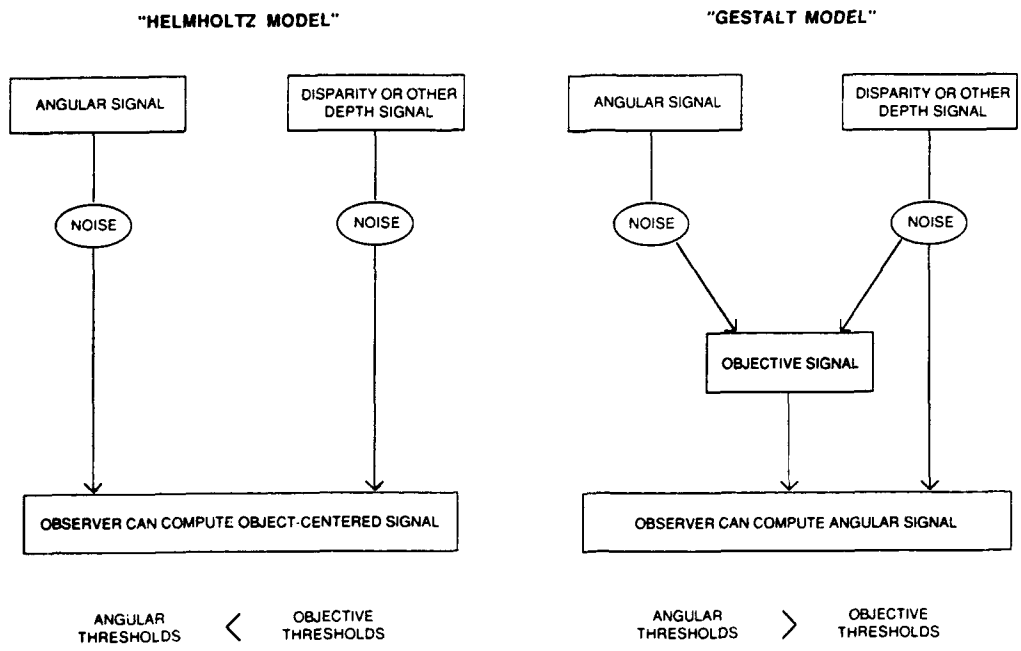


Fig. 1

signal based on relative depth. The observer must estimate angular dimensions indirectly by reversing the depth-correction. It seems likely that this reverse correction would degrade performance, and that generally the discrimination of objective dimensions would be superior to the discrimination of angular dimensions.

Burbeck (1987) used a discrimination approach to study "size" constancy for sinusoidal gratings. Surprisingly, her measurements with gratings presented at two different distances showed that observers can discriminate differences in objective spatial frequency (c/cm) as well or better than they discriminate differences in angular spatial frequency (c/deg). She concluded that human observers do not have direct access to the angular signal for spatial frequency, and perhaps not to any signal based on angular size (see also Gogel, 1969).

In our experiments we will compare the precision of angular and objective velocity discrimination in the presence of random trial-to-trial variations in target disparity. To determine whether size and velocity constancy share a common set of neural operations, we will use the same paradigm to measure angular and objective size discrimination.

METHODS

In any measurement of velocity discrimination, it is important to establish that the judgments are actually based on velocity and not on some co-varying dimension of the stimulus. For example, if the target is presented for a fixed duration, the observer can judge velocity from the distance traversed; if the targets cover a fixed distance, then velocity can be guessed from the target duration. To eliminate the consistent use of these confounding cues, we randomly varied the target duration by $\pm 20\%$ in most conditions; in a few conditions, the random variation in duration was increased to $\pm 30\%$.

We employed the method of single stimuli to avoid using a fixed "standard", since the standard velocity must necessarily by presented at a particular disparity. In the method of single stimuli, the observer is shown one of five velocities chosen from a narrow range, and is asked to indicate whether the presented sample is faster or slower than the mean of that range. No explicit standard or reference velocity is ever presented; instead the observer judges the sample against an implicit standard. A similar procedure was used for the size judgments, except that a larger number of sizes was included in the test set—13 sizes equal to each of the distances traversed by the moving target in the comparable experiment on velocity discrimination; the independent variations in speed and target duration resulted in this large number of distances. Experimental sessions began with a brief period of practice; 20 trials were usually

sufficient to define the mean and the range of the stimuli used for one experimental run. No feedback was given in the first experiment. Error feedback was given for the experiment which compared objective and angular velocity discrimination, and also for the experiment on objective and angular size discrimination.

The target for these experiments was a bright line, 13 min of arc in length and about 1 min arc in width, drawn by computer-generated signals on the screens of two Hewlett–Packard 1332A monitors, each equipped with a P-4 phosphor. The images on the two CRT screens were superimposed by a beam-splitting pellicle. Orthogonally-oriented polarizers placed in front of the CRT screens and the subject's eyes guaranteed that only one screen was visible to each eye. This arrangement allowed us to vary target disparity without changing target size or luminance. The target motion was never physically continuous, but was instead sampled at close spatial and temporal intervals—1.3 min arc/2.2 msec. As this sampling rate is close to the spatial and temporal resolution of the human visual system for these conditions, the motion appeared continuous. To measure the target luminance, we created a square patch 2.5 mm on a side, composed of 25 points equal in intensity to the test line (the diameter of each point was about 0.5 mm in size). The luminance measured with a Pritchard spectrophotometer through a proble which viewed about 4 points was found to be 3000 cd/m^2. The measured luminance of the dark background was about 0.8 cd/m^2. Ambient illumination, supplied by indirect fluorescent lighting, was at a moderate photopic level, so that equipment and furniture in the laboratory were easily visible.

In the first experiment, a vertical line was moved horizontally at a mean velocity of 9.9 deg/sec. In the subsequent experiments on size and velocity constancy, the target was a horizontal line moving vertically, also at a mean velocity of 9.9 deg/sec or alternatively 26 cm/sec. The direction of motion (left or right in the first experiment, up or down in subsequent experiments) was randomized from trial-to-trial in an effort to randomize the effects of anticipatory pursuit on target velocity (Kowler and Steinman, 1981). The mean target duration was 150 msec for subjects SM and LW; the mean duration was increased to 180 msec for subject NW. As indicated above, the duration was varied randomly from trial-to-trial. A fixation cross was presented prior to the appearance of the moving test line, adjacent to the mid-point of the traverse. The fixation cross remained visible during the presentation of the moving line, and for a short time after the test line disappeared.

We were concerned that the box housing the pellicle might serve as a "reference frame", so we repeated the experiment comparing velocity and size discrimination for objective and angular criteria in total darkness. The luminance of the target was reduced and the subject was first light-adapted so that even the faint glow from the CRT screen was invisible. The light-adaptation was refreshed every twenty trials. The pattern of results for both size and velocity discrimination was similar to the results obtained when the laboratory equipment was visible.

Each of the Weber fractions presented in this paper is based on at least 280 trials. A psychometric function was generated by plotting the percentage of trials on which the subject responded that the tested stimulus was faster (or larger) than the mean. A cumulative normal curve was fitted to the psychometric function by probit analysis. Threshold was defined as the incremental change that produced a change in the response rate from the 50% to the 75% level, equivalent to a d' of 0.675. Probit analysis also provided an estimate of the standard error of the threshold, which generally amounted to less than 10% of the estimated threshold.

Two additional experiments used different techniques to estimate the quality of the depth signal produced by disparity. In one experiment, the observer adjusted the distance of an external cardboard marker (a thin vertical rectangle with a pointed top) until it appeared to match the distance of the moving target visible inside the confines of the stereoscope. The observer was allowed to look back and forth until she was satisfied with the match. The laboratory walls, equipment and furniture surrounding the vertical marker were easily visible, and supplied abundant natural cues about the marker's physical distance. At least three of these matches were made for each of the test disparities presented on the CRT screens of the electronic stereoscope.

In a second experiment, observers were shown the same set of disparate moving targets used in the constancy experiments, and were asked to identify the target disparity by calling out a number between 0 and 9. Observers were given considerable practice labeling the disparities which were presented in random order.

Each observer then judged about 700 presentations; feedback about the correct disparity of the presented sample was supplied after each judgment.

Finally, we measured the influence of monocular cues to depth on velocity discrimination. For this experiment, the observer sat in front of a single CRT in well-illuminated surroundings and judged the velocity of a single moving point. In one condition, the observer rocked back and forth on every trial, so that her viewing distance changed from 28 to 57 cm on alternate trials. Obviously the range of angular velocities also changed by a factor of two on alternate trials. In the control condition, the observer viewed the target from a fixed distance (either 28 or 57 cm) during any given experimental session and the thresholds for the two distances were averaged. In both conditions, the observers viewed the screen monocularly; one eye was covered with a black eye patch. As before, the duration was randomly varied by ±20% from trial-to-trial obscuring distance and duration cues to velocity; the mean duration was 150 msec. The mean velocity was either 10 deg/sec at 28 cm or 5 deg/sec at 57 cm.

The two authors, and a third female observer who had never participated in velocity experiments before, served as the subjects. The targets were viewed with natural pupils at a distance of 1.5 meters except where indicated. The visual acuity of all three observers was 20/20 or better at this viewing distance.

RESULTS

Velocity discrimination with variable disparity

In the first experiment, we determined the basis of velocity discrimination in the presence of large, random trial-to-trial variations in disparity. Five different disparities were used covering a range of ±40 min arc. Because there was no explicit standard and no feedback, the observers were free to choose their own response criterion. They could have responded on the basis of either angular or objective velocity, but their responses were scored according to the angular velocity of the target.

If the observers had based their decisions on some estimate of *objective* velocity, then the random variations in disparity should have elevated their increment thresholds for *angular* velocity when compared to their thresholds for targets presented in a single plane. The Weber fractions on the right of Fig. 2, calculated from

VELOCITY DISCRIMINATION WITH RANDOM VARIATIONS IN DEPTH

	ΔV/V FIXATION PLANE 10 deg/sec	ΔV/V 5 MIXED DISPARITIES 10 deg/sec
LW	060 ± 006	063 ± 003
SM	040 ± 004	049 ± 002

Fig. 2. Velocity Weber fractions for a target presented only in the fixation plane are shown on the left. Velocity Weber fractions based on pooled data from targets presented at random with one of five disparities are shown on the right.

data pooled over all disparities, show that the disparity variations had almost no effect on the angular velocity threshold; they are nearly identical to the Weber fractions found for targets presented in the fixation plane alone. Figure 3 shows the thresholds for each of the separate disparities, and again there is no evidence of a systematic effect related to disparity. The observers made precise judgments of angular velocity despite random changes in target depth. This conclusion is reinforced by examining the medians (P.S.E.'s) of the individual psychometric functions for each disparity. If the observers were attempting to compensate for perceived distance—if they were responding to objective velocity—then there should be some evident shift in the median angular velocity associated with each disparity. None is apparent in the data shown in Fig. 4.

These results do not, by themselves, prove that there is no velocity constancy. They show only that well-practiced observers can choose to base their judgments on the angular velocity signal, and that this signal is unaffected by manipulations of target disparity.

Size and velocity discrimination for angular and objective dimensions

It is a common observation in research on size constancy that observers will match either the

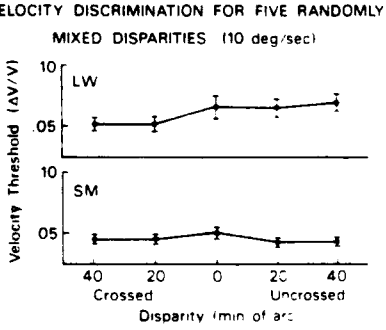


Fig. 3. The individual velocity Weber fractions for the five disparities, presented at random in the experiment.

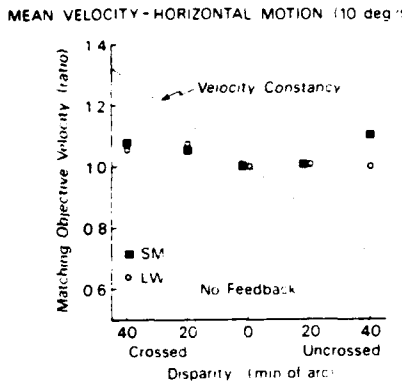


Fig. 4. The medians of the separate velocity psychometric functions (P.S.E.'s) for each of the five disparities. There is no systematic shift in the velocity P.S.E. with changes in disparity. No feedback was supplied in this experiment.

WEBER FRACTIONS

VELOCITY ($\Delta V/V$)

	ONE PLANE	ANGULAR VELOCITY	OBJECTIVE VELOCITY
LW	039 ± 004	052 ± 006	118 ± 016
SM	049 ± 005	053 ± 006	093 ± 011
NW	086 ± 006	089 ± 006	142 ± 009
AVE	06	06	12

SIZE ($\Delta S/S$)

	ONE PLANE	ANGULAR SIZE	OBJECTIVE SIZE
LW	045 ± 005	053 ± 005	063 ± 007
SM	044 ± 004	065 ± 007	068 ± 006
NW	062 ± 005	092 ± 007	106 ± 008
AVE	05	07	08

Fig. 5. A comparison of size and velocity discrimination. The first column shows the Weber fractions for targets presented in a single plane. The second column shows the Weber fractions when feedback reinforced judgments based on angular velocity or angular size. The third column shows the Weber fractions when feedback reinforced judgments based on objective velocity or size.

angular or objective size of a target depending on the instructions of the experimenter (Gilinsky, 1955; Leibowitz and Harvey, 1969; Carlson, 1977). No instructions or feedback were supplied in the first experiment, but the two authors who served as subjects could have made an unconscious choice in favor of angular velocity. We therefore repeated the first experiment with explicit feedback reinforcing choices based either on angular or objective target velocity in separate experimental sessions. We also used 10, instead of five, disparities covering a ± 45 min range, and randomly interspersed them on successive trials, but with a bias in favor of the disparities farthest from the fixation plane in order to enhance the difference between objective and angular velocity.

Velocity constancy might be just a simple extension of size constancy (Rock *et al.*, 1968; Epstein, 1973). An observer could judge objective velocity by judging the time the target took to traverse an *objective* distance estimated through a size constancy mechanism. How precisely can human observers judge the size of the traverse? Are they better at judging angular size than objective size? In a companion study to the velocity discrimination measurements, we asked our observers to judge whether the length of the traverse was longer or shorter than the mean distance covered by the moving target.

The Weber fractions for the six different conditions are shown in Fig. 5. These data were taken after the observer had had considerable practice in all experimental conditions. The first column gives the Weber fractions for the control condition in which all targets are presented in the fixation plane. As in the previous experiment, judgments of angular velocity are not affected by random variations in disparity (second column), but error feedback is not sufficient to produce good objective velocity discrimination (third column). The Weber fractions for objective velocity discrimination are significantly higher than those for angular velocity discrimination.

The results for size discrimination are quite different. The random variations in disparity elevate the Weber fractions somewhat when compared to size discrimination in a single plane, but it does not matter whether the observer is trying to judge objective or angular size. There is no significant difference between angular and objective size discrimination for any of the three observers. Size constancy is decisively more robust than velocity constancy.

Perceived distance and disparity

In our displays, the only cue to distance was target disparity: there were no changes in physical size, brightness or other ordinary cues to distance. We wondered if the disparity information were sufficient to provide veridical information about physical distance. Previous studies had shown that differences in disparity supply adequate information about differences in physical distances for viewing distances ranging from 1 to 4 m (Foley, 1980). Our viewing distance was 1.5 m. Nevertheless, the conditions here—a briefly-presented moving target—may not have been ideal for using relative disparity to estimate relative distance.

To determine the perceived distances associated with targets viewed in the electronic stereo-

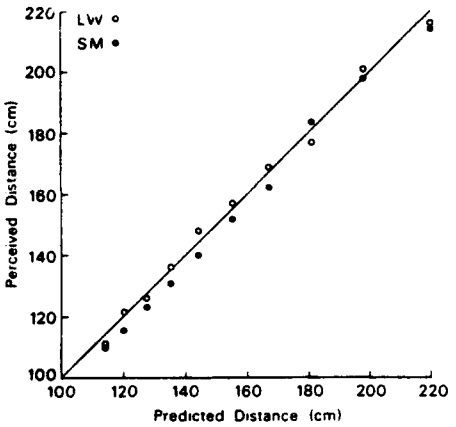


Fig. 6. Observers were asked to match the distance of an external marker to the apparent depth of a moving disparate target presented in the stereoscope. The ordinate is the matching distance of the external marker, and the abscissa is the predicted distance calculated from the relative disparity of targets presented at 1.5 m.

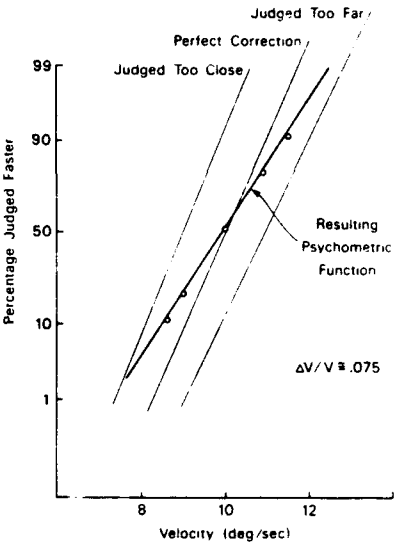


Fig. 7. A diagram of the correction needed to produce precise objective velocity discrimination. See explanation in text.

scope, we asked observers to place an external marker at a distance matching the perceived distance of each of the moving targets. The ordinate of Fig. 6 corresponds to the physical location of the marker at the matching distance. The abscissa is the predicted distance of the target in the stereoscope, calculated by assuming a straight-forward transformation of disparity *into* distance, relative to a viewing distance of 1.5 m.* Clearly, perceived distance and predicted distance agree remarkably well. The observers may have merely matched the disparities of the two stimuli, but the enhanced distance cues associated with the exterior marker did not disturb this disparity match.

Taking account of the distance error

We next estimated the variability (or error) associated with the estimate of distance. The disparity steps were very large and thus easily discriminable, but constancy correction requires more than discrimination between simultaneously presented disparities. To estimate the objective dimensions accurately, our observers had to identify, at least implicitly, which of the 10 disparities (or distances) was presented on each trial. The precision of their ability to identify different disparities was estimated directly by asking them to label the disparities. On average, the observers identified each disparity correctly on about one-third of the presentations, and were seldom off by more than one step on either side of the correct disparity.

Could we predict objective velocity discrimination by combining the error in angular velocity discrimination with the error in estimating the physical distance? Let us assume that the observer takes the angular velocity signal and corrects it by the disparity signal. This operation is equivalent to shifting the psychometric function for angular velocity from the actual plane of the target towards the fixation plane. A diagram of the operation is shown in Fig. 7.† Consider what will happen to a target presented far behind the fixation plane, so that the mean angular velocity is about 8 deg/sec. If the observer makes a perfect correction, the function will be shifted to a position centered on the mean velocity of 10 deg/sec (see line labeled "perfect correction"). If the observer assumes that the target is one disparity step farther from the fixation plane than is actually the case, she will shift the psychometric function too much, as shown by the line labeled "judged too far". An under-correction results from assuming that the target is one step closer than is actually the case ("judged too close"). If we assume that

*We used the simple approximation,

$$s = a/\tan(D + \Delta d)$$

where s is the distance in centimeters, D is the convergence angle for a viewing distance of 1.5 m and equals 2.29 deg, d is the incremental change in the angle associated with each tested disparity, and a is the interpupillary separation assumed to be 6 cm.

†The psychometric functions are plotted on probability paper which transforms a cumulative normal curve into a straight line.

WEBER FRACTIONS FOR OBJECTIVE VELOCITY

	PREDICTED	OBSERVED
LW	062	118 ± 010
SM	071	093 ± 011
NW	097	142 ± 009

Fig. 8. The first column shows the predicted Weber fractions for objective velocity discrimination based on the angular velocity Weber fraction and the errors in disparity judgments. The second column shows the measured Weber fractions for objective velocity discrimination.

each of these events happens on one third of the trials, then the measured psychometric function will look like the darker line superimposed on the other three. The resulting psychometric function is necessarily shallower because the observer is effectively averaging his responses from three different psychometric functions. Therefore, the measured Weber fraction for objective velocity is about 7.5%, or 1.5 times the Weber fraction for angular velocity which, for this example, was assumed to be 5%.

The complete calculation was more complicated than the diagrammed example. We had presented more targets at the extreme disparities than near the fixation plane, and some of the disparities were more difficult to identify than others. Moreover, both the psychometric function and the pattern of disparity errors varied from observer to observer. To predict the exact elevation in the objective velocity threshold when compared to the angular velocity threshold, we used each observer's psychometric function for angular velocity, and weighted it by their particular pattern of errors in judging each test disparity. That is, we assumed that the observers were adjusting the velocity signal for depth with the accuracy indicated by their ability to label each disparity.

The results of this calculation are shown in the first column of Fig. 8; the measured Weber fractions for objective velocity discrimination are in the adjacent column. The predicted values are smaller than observed values; the disparity error would have to be larger than our data indicate to account for the observed error in objective velocity discrimination. Human observers are inefficient in combining these two dimensions to achieve constancy.

How then did our observers judge objective velocity? To prevent our observers from making consistent use of duration alone as a cue to velocity, we randomized the duration from trial-to-trial, but they still could judge velocity by separately estimating both the duration and the total distance traversed on *each* trial. Angular velocity discrimination is too precise to be judged on this basis, but objective velocity might be judged by combining an estimate of duration and an estimate of *objective* size (Rock *et al.*, 1968; Epstein, 1973). The Weber fraction for duration is about 0.10 (McKee and Taylor, 1984) and our measured Weber fraction for objective size discrimination is 0.08. A simple combination of these errors would predict a Weber fraction for objective velocity of 0.13 which is close to what we observe:

$$\sqrt{(\text{objective size error})^2 + (\text{duration error})^2} = \text{objective velocity error:}$$

$$\sqrt{(0.08)^2 + (0.10)^2} = 0.13.$$

Size and duration variables do influence velocity matching. Using targets viewed in total darkness, Rock *et al.* (1968) found that observers took account of disparity in matching the apparent velocities of near and distant targets, but this tendency to constancy was substantially weakened when the durations of the near and distant targets were matched. Wist *et al.* (1976) used the Pulfrich effect to manipulate the apparent depth of moving gratings, and found some tendency towards velocity constancy, but they also reported that changes in spatial frequency had a far more significant effect on the perceived velocity.

Velocity transposition and velocity discrimination

The results described so far suggest that velocity constancy does not exist. There is no efficient compensation for depth or distance embedded in the neural machinery used to estimate velocity. Yet this conclusion contradicts common experience. Display a drifting grating on a CRT and walk towards the screen: the apparent speed of the grating does not change in any obvious way as you approach the display.

Many years ago, Brown (1931) observed a phenomenon known as velocity transposition—apparent velocity depends on the ratio of the angular velocity to the angular size of the surrounding frame or aperture. Brown found that his observers would even accept a match between two velocities differing by a factor of two, and presented at the same distance, if the slower moving target was surrounded by a frame of half the size of the frame surrounding the faster target (see also Gogel and McNulty,

"FRAME" EFFECT

WEBER FRACTIONS (ΔV/V)

	SINGLE PLANE	MIXED PLANES
LW	.071 ± .006	.139 ± .031
SM	.057 ± .005	.087 ± .012
AVE	.064	.113

Fig. 9. The left hand column gives the Weber fractions for targets viewed from a fixed distance during a given experimental session; values are the average of data from two distances (28 and 57 cm) with corresponding changes in the mean angular velocity (10 and 5 deg/sec). The Weber fractions on the right were obtained when the observer changed her distance from 28 to 57 cm on alternating trials, producing concomitant trial-to-trial changes in the mean angular velocity. Both sets of data based on monocular viewing of a well-illuminated CRT screen.

1983). In his brilliant essay on velocity constancy, Wallach (1976) concluded that this transposition effect was sufficient to account for constancy, because as one approached a moving object, the natural surroundings framing the target would necessarily increase in angular size preserving the ratio between the angular velocity and the surroundings; no compensatory mechanism based on the depth signal was required. But how does transposition work? It is conceivable that some automatic neural process calculates the ratio of angular velocity to angular size in order to estimate objective velocity, but there is a simpler explanation. The observers may be matching the time each target takes to move across its respective frame (Smith and Sherlock, 1957).

We wondered how well transposition ("frame" effects) would fare in our paradigm, where the target does not traverse the whole frame of the CRT, and target duration is varied randomly from trial-to-trial. Observers made monocular judgments of velocity while changing their distance from the target on every trial. The error feedback was the same as if the observers were stationary, i.e. they were rewarded for compensating for the change in viewing distance. If they were able to compensate perfectly for the change in distance, their performance would equal the velocity discrimination for targets viewed at a single distance.

The column in Fig. 9 entitled "mixed planes" shows the Weber fractions for the changing distance condition. Despite the random variations in duration, the observers were able to use the change in overall scale to compensate for the changes in distance. Transposition works as well as, or perhaps slightly better than, disparity compensation in producing velocity constancy, but again the transposed velocity judgments are less precise than angular velocity discrimination for a single plane.*

*For technical reasons, we were unable to measure angular velocity discrimination for the mixed planes condition. We had no means of sensing the position of the observer so that we could program a compensating change in angular velocity with changes in distance.

DISCUSSION

Constancy is thought to be an automatic process which effortlessly translates the angular dimensions of the retinal stimulus into an object-centered code. By that criterion, there is no velocity constancy. Human observers can make imprecise judgments of objective velocity, but their performance is inefficient, particularly when compared to their ability to judge objective size. Instead the human visual system preserves a precise angular velocity signal uncorrected for depth.

What good is a velocity signal which cannot be used to encode the speed of real objects? Velocity information serves many other functions in visual processing. It can be used to encode depth, define object boundaries, and direct eye and body movements (Nakayama, 1985). All of these functions depend on a velocity signal encoded in *angular* units; for example, the sensory input for oculomotor smooth pursuit is angular velocity. The sensory input to motor centers is undoubtedly processed in parallel to the perceptual input. The angular velocity signal may be sent directly from the retina to the oculomotor control centers, as well as to other motor centers, and the perception of velocity could be based on a different set of neural operations.

A more complicated problem concerns the perceptual use of the angular velocity signal. One putative constancy mechanism, velocity transposition, depends on the perceived ratio between angular velocity and angular size. As a mechanism for constancy, velocity transposition has one significant virtue—apparent velocity does not change with changes in the observer's position, and this helps preserve object identity during perspective changes. Transposition also suffers from a major defect—accurate judgments of objective velocity depend on a fortuitous arrangement between the moving target and its surroundings (or "frame").

Consider two objects at different physical distances from the observer, moving at the same physical velocity, against a natural background consisting of a fairly regular texture (grass, pavement, rocky surface). The receding textural gradient could set the angular size scaling for velocity transposition thereby producing the veridical percept that the physical velocities of the two objects are the same. There is, however, a competing tendency due to size constancy which may partially cancel the angular information contained in the textural gradient. Again we can invoke parallel processing to allow both constancies to function, but suppose the background is irregular, so that one of the moving objects is framed by a local aperture, e.g. a gap in the trees. Now the transposition mechanism for velocity constancy will necessarily fail, as it did in Brown's experiment (see Results section).

Perhaps our most significant perceptual use for angular velocity is to estimate physical distance via motion parallax. When we translate through our environment, distant objects move more slowly across our retinae than near objects. Usually we have co-varying disparity information about the distance to our surroundings. If we used the disparity signal to correct the velocity signal as constancy compensation requires, then we would lose the motion parallax information for relative distance. Motion parallax information could also be processed in parallel, permitting the observer to perceive a veridical velocity signal, but our results suggest this does not happen. There is no partitioning of perceptual experience which allows us to perceive objective velocity and motion parallax simultaneously.

It could be argued that we would have found evidence for velocity constancy had we used more complex, natural targets, although it is difficult to understand why velocity constancy would require a more complex target than size constancy. To answer this objection, we ask that you perform a simple armchair experiment. Look across a cluttered room and fixate on some distant object; move from side to side so that the objects in the room translate across your retinae. Do the nearest objects appear to move faster than the farthest ones? This simple demonstration of motion parallax is evidence that the velocity signal is not corrected by the signal for depth even in complex environments.

Acknowledgements—We want to thank Dr Eileen Kowler for helpful discussions on oculomotor smooth pursuit. We also want to thank Dr Samuel Bowne for his perceptive and critical comments on earlier versions of this manuscript, and Dr Ken Nakayama and Dr Robert Steinman for suggesting the experiment on velocity transposition. Nance Wilson served as a diligent and patient "naive" subject for these experiments. This work was supported by USAF grant AFOSR-85-0380, and NIH core grant 5P-30-EY-01186.

REFERENCES

Brown J. F. (1931) The visual perception of velocity. *Psychol. Forsch.* **14**, 199–232.

Burbeck C. A. (1987) On the locus of spatial frequency discrimination. *J. opt. Soc. Am. A.* **4**, 1807–1813.

Carlson V. R. (1977) Instructions and perceptual constancy judgments. In *Stability and Constancy in Visual Perception: Mechanisms and Processes* (Edited by Epstein W.), pp. 217–254. Wiley, New York.

Epstein W. (1973) The process of "taking-into-account" in visual perception. *Perception* **2**, 267–285.

Foley J. M. (1980) Binocular distance perception. *Psychol. Rev.* **87**, 411–434.

Gilinsky A. S. (1955) The effect of attitude upon the perception of size. *Am. J. Psychol.* **68**, 173–192.

Gogel W. C. (1969) The sensing of retinal size. *Vision Res.* **9**, 1079–1094.

Gogel W. C. and McNulty P. (1983) Perceived velocity as a function of reference mark density. *Scand. J. Psychol.* **24**, 257–265.

Helmholtz H. V. (1868) The recent progress of the theory of vision. Translated by Pye-Smith P. H. In *Popular Scientific Lectures* (edited by Kline M.) Dover Publications, New York (1982).

Kowler E. and Steinman R. M. (1981) The effect of expectations on slow oculomotor control—III. Guessing unpredictable target displacements. *Vision Res.* **21**, 191–204.

Leibowitz H. W. and Harvey L. W. (1969) Effect of instructions, environment, and type of test object on matched size. *J. exp. Psychol.* **81**, 36–43.

McKee S. P. and Taylor D. G. (1984) The discrimination of time: a comparison of foveal and peripheral sensitivity. *J. opt. Soc. Am.* **1**, 620–627.

Nakayama K. (1985) Biological image motion processing: a review. *Vision Res.* **25**, 625–660.

Rock I., Hill A. L. and Fineman M. (1968) Speed constancy as a function of size constancy. *Percept. Psychophys.* **4**, 37–40.

Smith O. W. and Sherlock L. (1957) A new explanation of the velocity-transposition phenomenon. *Am. J. Psychol.* **70**, 102–105.

Wallach H. (1976) On the constancy of visual speed. *Psychol. Rev.* **46**, 541–552 (1939). Reprinted in *On Perception*, pp. 149–160. Quadrangle/New York Times Books.

Wist E. R., Diener H. C. and Dichgans J. (1976) Motion constancy dependent upon perceived distance and the spatial frequency of the stimulus pattern. *Percept. Psychophys.* **19**, 485–491.

Woodworth R. S. (1938) *Experimental Psychology*. Holt, New York.

SUZANNE P. MCKEE

Neural coding of local and global motion

Perceptual and electrophysiological evidence has been found for the integration of local motion signals into a single global motion.

When we view a complex object in motion, we have relatively little trouble assigning it a unique direction and speed, even though many of its components may not move with either the speed or direction of the object as a whole. Consider our view of a horse galloping down the straightaway. The legs move rhythmically to and fro, the head rocks at an odd angle to the forward direction and, as anyone who has ridden a horse can attest, the back is moving predominantly up and down. Yet somehow all these diverse motions are integrated into an image of the whole beast moving straight down the track.

Visual scientists generally use simple 'nonsense' images formed of points or lines to study how the local motions of individual features or components are blended into a global motion percept. A commonly used target consists of numerous bright dots randomly distributed on the flat screen of a computer monitor (see Fig. 1). The dots are flashed briefly in one set of positions and then, a few milliseconds later, are flashed again in new positions; each flashed presentation is like a 'frame' in a movie. If the position of each dot is changed incrementally in each subsequent frame by the same small distance and in the same direction, then the dots will appear to move smoothly in one direction (Fig. 1a). If, on the other hand, the direction of the change in position is randomly selected for each dot, then the dots will appear to move incoherently, like 'snow' on a television screen (Fig. 1b).

Random dot displays have been used to determine the range of local directions that can be integrated into the global percept of motion in a single direction [1]. The range of directions that each dot could move from frame to frame were systematically varied around vertical, and observers were asked to judge whether the overall display appeared to be flowing upward.

Surprisingly, the percept of upward flow was seen consistently as long as the range of individual dot directions did not exceed 180° (Fig. 1c), provided that the size of the spatial step taken by each dot in each frame was less than one degree in visual angle. Perhaps our visual system takes the average of the directions and reports 'upward', if no sizeable number of dots are actually moving downward. In fact, the process is more elaborate than that. Using an overall range for dot directions of 180°, Williams and Sekuler showed that, if the distribution was constrained so that no dots moved in a direction close to the mean direction (that is, within ± 20° of vertical), the percept of upward flow diminished sharply [1]. This shows that some local signals moving in the global direction are needed to produce a unified percept.

The neural computation that underlies the percept of global flow is remarkably precise. If all the dots in the display move coherently as in Fig. 1a, observers can accurately detect a 1° angular shift in direction. This same 1° precision is found when the local directions taken by the individual dots in the display span a range of 30° centered on the mean [2]. In fact, when the range of local directions is increased to over 100°, the global direction can still be judged to a precision of 3 or 4°.

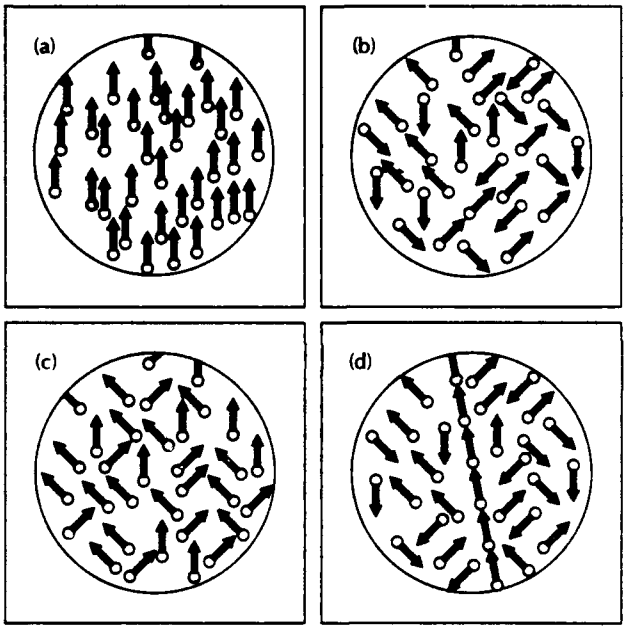


Fig. 1. Diagram of dynamic random dot displays. (a) All dots hop in the same direction and by the same distance, producing the appearance of coherent upward motion. (b) All dots hop in randomly chosen directions (range of directions, 360°), producing appearance of motion jitter. (c) All dots hop in randomly chosen directions with range limited to 180°, producing appearance of global upward flow. (d) Detecting a local trajectory: several frames of dot motion in which a single dot always hops in the same direction and the remainder of the dots hop in randomly chosen directions.

Watamaniuk *et al.* explained the precision of direction discrimination by assuming the existence of cortical motion mechanisms broadly tuned to twelve different directions spanning the compass [2]. Each hypothetical mechanism consists of a group of neurons that have identical directional tuning and differ from each other only in spatial position. It is assumed, for example, that neurons are available to sense vertical motion everywhere in the field of view, but that each neuron is responsive to a dif-

ferent portion of the visual field. Imagine a set of neurons from different mechanisms, hence with different directional sensitivity, receiving the signals from the retina produced by viewing the display shown in Fig. 1c. Each tuned neuron detects all the dots going in its preferred direction, but because its tuning is broad, overlapping the directional tuning of neurons from other mechanisms, it also responds somewhat to other directions. Presumably, we see a unique global direction when the neural mechanism tuned to vertical has a strong signal, and the signals in mechanisms tuned to off-vertical directions are in rough balance. Our precision in detecting small shifts in the global direction is thought to depend on the ratio of signals in these off-vertical directional mechanisms.

Even when a global direction is perceived, the directions moved by the individual dots are still apparent. Whether local or global motion dominates the percept depends to some extent on voluntary shifts in attention. Ross (personal communication) has demonstrated that a single dot following a straight trajectory can be detected in the midst of dots in random incoherent motion (Fig. 1d). In fact, this trajectory point is easily detected even in very dense random dot displays [3]. As all the dots are identical (the trajectory point is shown with a dark red arrow in the diagram for clarity), and each dot hops to a new location on each frame, how does the visual system identify the dot hopping along a fixed trajectory from all the other moving dots? One might expect that numerous mismatches between the trajectory dot and the background dots would obscure its trajectory. In computational studies for machine vision, this matching problem is known as the 'correspondence problem'. Recently, Grzywacz, Smith and Yuille [4] described a computer algorithm, which they called 'temporal coherence', that enforces directional consistency over time to solve the correspondence problem. In their computer simulation, dots were constrained to make the match in the present frame that minimized the change in direction from the match made in the previous frame — a constraint that increased the speed and accuracy of their algorithm compared to alternative formulations. Presumably, something like 'temporal coherence' is embedded in the functional structure of the neurons that form the directional mechanisms described above.

Which neurons in the visual cortex are responsible for integrating the local signals into a unique global direction? Many neurons in primary striate cortex (V1) are selective for direction of motion, but they are also strongly sensitive to target contrast and, for bar or line targets, to target orientation. Thus, the responses of the V1 neurons confound changes in direction with changes in contrast or orientation. For the past decade, considerable interest has been focused on a special extrastriate visual area, the middle temporal (MT) area, which receives substantial input from the directionally selective units in V1 [5]. MT lies in the posterior bank of the superior temporal sulcus, and most of the neurons in this area respond selectively to the direction and speed of moving targets [6]. Now, their response to target contrast has been shown to saturate at relatively low contrast levels, so changes in target contrast are not confounded with changes in target motion [7]. Moreover, neurons in MT are responsive to stimuli falling within much larger segments of the field of view than V1 neurons, making them ideal candidates for the type of large-scale visual integration needed in the calculation of global motion [5,7,8].

Following the experimental destruction of a portion of area MT by injections of ibotenic acid [9], sizeable decrements were found in the ability of macaque monkeys to discriminate upward from downward motion in partially coherent random dot targets; coherence was varied by increasing the percentage of dots moving either upward or downward in an otherwise incoherent target. There were no losses in the ability of these monkeys to make contrast judgments. Similarly, Pasternak *et al.* [10] noted substantial deficits in the precision of direction discrimination for global motion after they removed MT plus an adjacent region in the superior temporal sulcus, MST, from the cortex of macaque monkeys. However, the discrimination of local directional signals (the direction of drifting bar targets) was not impaired.

These results strongly suggest that the capacity to identify the global direction specified by the integrated local directional signals depends on the integrity of the MT/MST region of extrastriate cortex of the macaque, and its homologue in human cortex.

References

1. WILLIAMS DW, SEKULER R: **Coherent global motion percepts from stochastic local motions.** *Vision Res* 1984, 24:55–62.
2. WATAMANIUK SNJ, SEKULER R, WILLIAMS DW: **Direction perception in complex dynamic displays: the integration of direction information.** *Vision Res* 1989, 29:47–59.
3. MCKEE SP, WATAMANIUK SNJ: **Detecting a single point moving on a linear trajectory amidst randomly moving points.** *Invest Opbthalmol Vis Sci* 1991, 32 (suppl), in press.
4. GRZYWACZ NM, SMITH JA, YUILLE AL: **A common theoretical framework for visual motion's spatial and temporal coherence.** *Proc Workshop Vis Motion, IEEE* 1989, 148–155.
5. MAUNSELL JHR, VAN ESSEN DC: **Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation.** *J Neurophysiol* 1983, 49:1127–1147.
6. MAUNSELL JHR, NEWSOME WT: **Visual processing in monkey extrastriate cortex.** *Ann Rev Neurosci* 1987, 10:363–401.
7. SCLAR G, MAUNSELL JHR, LENNIE P: **Coding of image contrast in central visual pathways of the macaque monkey.** *Vision Res* 1990, 32:1–10.
8. WATAMANIUK SNJ: **Information loss in the integration of direction information.** *Invest Ophthalmol Vis Sci* 1990, 31:519.
9. NEWSOME WT, PARÉ EB: **A selective impairment of motion perception following lesions of the middle temporal visual area (MT).** *J Neurosci* 1988, 8:2201–2211.
10. PASTERNAK T, MAUNSELL J, POLASHENSKI W, MERIGAN W: **Deficits in global motion perception after MT/MST lesions in a macaque.** *Invest Ophthalmol Vis Sci* 1991, 32, in press.

Suzanne P. McKee, Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, USA.

17 The Physical Constraints on Visual Hyperacuity

Suzanne P. McKee

Introduction

Under ideal conditions, human observers can judge the relative position of two visual features with a precision that is substantially smaller than the size of a foveal cone. Westheimer (1975) called these highly precise spatial judgments, hyperacuities, to distinguish them from measures of visual acuity that assess resolution capability. Almost any type of target configuration can be used to measure hyperacuity, but the most common varieties are shown in Fig. 17.1. For each type of target, the observer is asked to discriminate between two possible configurations, e.g. offset left or offset right in Vernier judgments. The smallest spatial difference that is correctly identified according to a particular statistical criterion is termed the hyperacuity threshold. In the central fovea of normal observers, hyperacuity thresholds are ordinarily less than 10″, the exact value depending on the selected statistical criterion and on numerous physical constraints. Here, hyperacuity is defined as any judgment of relative spatial position that is substantially better than resolution acuity, so that the term can be extended to spatial judgments made in the peripheral visual field as well as in the fovea.

Hyperacuity is not just a curiosity. The features shown in Fig. 17.1 represent the basic elements of any visual pattern, so hyperacuity thresholds can be considered simple, but rigorous, measures of pattern discrimination (Morgan, 1986). Do all these different spatial judgments depend on the same neural substrate? Probably not. Stereoacuity seems to depend on a special disparity mechanism that is quite different from the mechanism responsible for judgments of relative lateral position. For one thing, disparity thresholds can be significantly smaller than the thresholds for Vernier offset, spatial interval or bisection even when measured with an identical configuration (Berry, 1948; Westheimer and McKee, 1979; Schor and Badcock, 1985; McKee *et al.*, 1990). A similar argument can be made for a special mechanism for motion hyperacuity. If a target is abruptly displaced, the mini-

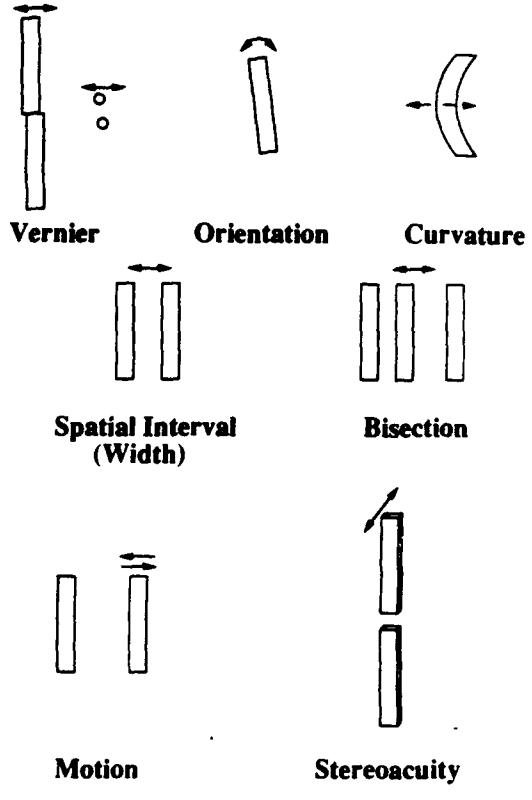


Fig. 17.1 *The hyperacuities: target configurations that can produce positional thresholds substantially smaller than resolution acuity. For the Vernier target, observer judges whether upper line or upper dot lies to the left or right of lower line or dot. Orientation: observer judges whether line is tilted left or right. Curvature: observer judges whether line is more or less curved than some mean value. Spatial interval: observer judges whether separation is larger or smaller than some mean value. Bisection: observer judges whether central test line is closer to left or right hand reference lines. Motion: observer judges whether an abruptly displaced line moved left or right. Stereoacuity: observer judges whether upper line is in front or behind lower line.*

mum displacement required to discriminate the direction of motion can be much smaller than the incremental difference in distance required to judge that one static spatial interval is larger than another, again for identical target configurations (Westheimer, 1979; Legge and Campbell, 1981; Nakayama and Tyler, 1981). There are also numerous reasons for suspecting that the other spatial judgments – orientation, curvature, Vernier offset, spatial interval – each depend on subtle differences in visual processing (Andrews *et al.*, 1973; Westheimer and McKee, 1977b; Watt, 1984). It is, therefore, difficult to generalize about the effect of any particular physical variable on 'hyperacuity' because there are, in fact, many hyperacuities, each with somewhat different stimulus requirements. Where possible, I have included graphs for many different hyperacuities in order to capture this diversity.

Luminance

Like visual resolution, the hyperacuities depend on an adequate amount of light for best performance. But is this dependence the same as visual acuity? In a series of thoughtful papers, Geisler analysed the theoretical improvement in two-point resolution and in two-point spatial interval judgments expected with increasing light (Geisler, 1984, 1989; Geisler and Davila, 1985). For these ideal observer calculations, Geisler assumed that performance was limited only by fluctuations in the number of quanta absorbed by the foveal cones. Not surprisingly, his calculations showed that spatial interval thresholds should fall as the reciprocal of the square root of the number of quanta absorbed. Any point source can be localized with a precision that depends on the square root of the number of quanta in the image. So whether one thinks of a spatial interval judgment as localizing two point sources and then estimating the distance between them, or perhaps more reasonably, as discriminating between two possible configurations, e.g. two points separated by either 3′ or 3.1′, the decreasing variance of the light distributions should produce an improvement of this type.

The interesting result of Geisler's calculations was the prediction that two-point resolution should improve as the reciprocal of the fourth power of the number of quanta absorbed. The difference between the predictions for resolution and hyperacuity undoubtedly arises because of the substantial overlap between the two point sources in the acuity task. As the one-dimensional example given in Fig. 17.2 roughly shows, the relationship between the separation of the points and the light absorbed in the receptors is not simple for overlapping distributions. For example, doubling the separation ((b) to (c)) produces about a 40% change in the quanta absorbed by the central cones and shifting from no separation (a) to a separation equal to distance between the receptors (b) produces an even smaller change. Geisler notes that visual resolution depends on information contained in the second derivative of the stimulus shape, while hyperacuity judgments (separation > 2′) depend on information in the first derivative.

Unlike the ideal observer, the real observer is unable to take full advantage of the increasing number of quanta in the stimulus. As has been known for many decades, resolution acuity is virtually independent of luminance above about 30 ml (Riggs, 1965). Geisler's own data are displayed in Fig. 17.3 where the resolution acuities for three observers reach asymptotic values at a moderate photopic level; the solid lines show Geisler's predictions for the ideal observer. Geisler attributes this deviation from prediction to Weber's law, a compressive non-linearity that reduces both quantal signal and noise prior to some post-receptoral source of noise.

Geisler's data for two-point spatial interval judgments also do not conform to the square root improvement in hyperacuity predicted for the ideal observer, but, curiously, they do appear to follow the fourth root improvement predicted for acuity. Probably these hyperacuity thresholds are merely approaching asymptotic performance, creating the illusion of a fourth root dependence. Nevertheless, this fourth root relationship has been observed in hyperacuity measurements made with quite different targets (Fig. 17.4). Kiorpes and Movshon (1990) noted that Vernier acuity improved as the fourth root of luminance for offset square-wave grating targets. The Vernier thresh-

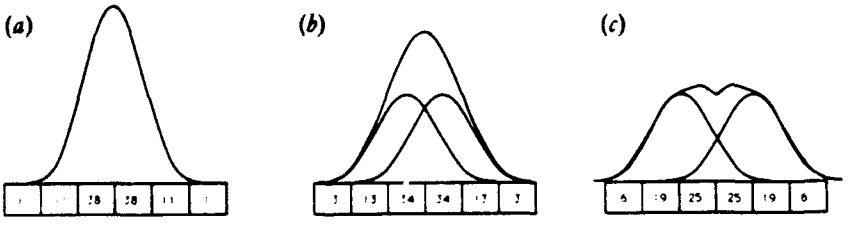


Fig. 17.2 *Two-point resolution: a rough one-dimensional example that shows the proportional change in number of quanta absorbed in receptors as separation between two point sources increases. Gaussian shape assumed for light distribution. (a) Two points superimposed (b) Separation equal to single receptor. (c) Separation twice that in (b).*

olds of Westheimer and Pettet (1990) also roughly follow this fourth root relationship as do the bisection measurements of Klein and Levi (1985). It is possible that the significant hyperacuity cue in these targets lies in the residual overlap between the light distributions associated with the lines or bars – a low luminance region that might resemble the arrangement for acuity measurements at low luminances. Note that bisection thresholds show this dependence on luminance only for the closely spaced lines used in the Klein–Levi study; thresholds for widely spaced target lines are independent of luminance except at the lowest value. Still there is no evidence in Geisler's data that acuity measurements ever show this fourth root dependence even at low luminances.

The luminance units used for the graph of the Klein–Levi results are somewhat unusual. It is difficult to specify the luminance of the narrow bars and lines in the oscilloscope displays commonly used in hyperacuity studies. Rather than the traditional squared units associated with extended sources, Klein and Levi (1985) and also

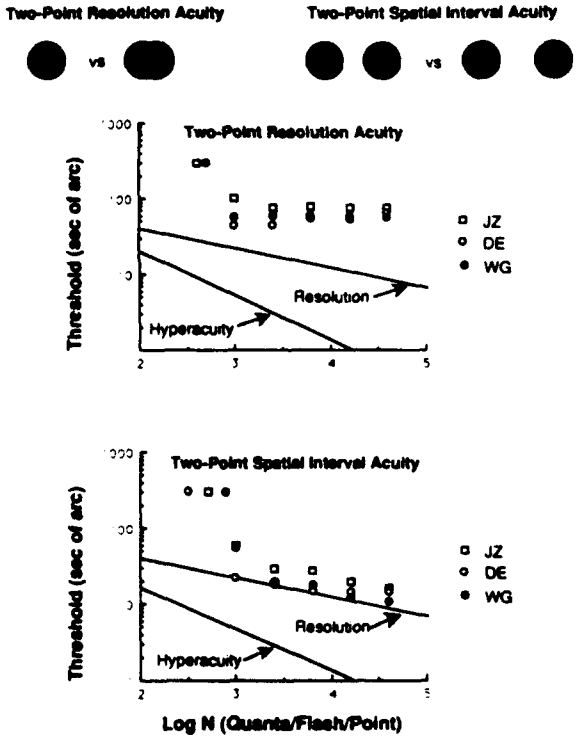


Fig. 17.3 *Data replotted from Geisler and Davila, 1985. Continuous lines show Geisler's predictions for two-point resolution (slope = −0.25) and for two-point spatial interval judgments (slope = −0.5). Target configurations are shown at the top. Points 0.2′ in diameter. Target duration 100 ms. Mean separation for spatial interval judgments was 3′. Foveal presentation.*

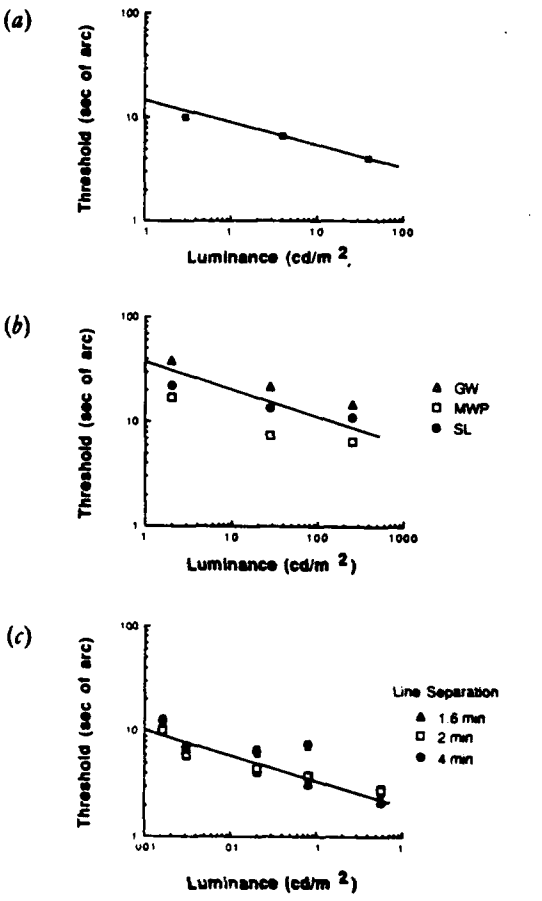


Fig. 17.4 *The effect of luminance on hyperacuity judgments. (a) Vernier acuity for sinusoidal grating (4.6 c deg^{-1}) replotted from data from Kiorpes and Movshon, 1990). (b) Vernier edge acuity for vertically aligned blocks at 17% contrast replotted from data from Westheimer and Pettet, 1990. (c) Bisection acuity for long thin lines for three mean separations between lines. (Replotted from data in Klein and Levi, 1985.)*

Westheimer (1985) recommend a linear measure (cd m^{-1}) more appropriate to line targets. Similar difficulties are encountered in specifying the contrast of thin lines or bars. For practical advice on measuring luminance, contrast, and the light distributions of oscilloscopic displays, the reader is directed to articles by Sperling (1971), Watt and Morgan (1982) and Westheimer (1985).

Contrast

Visual scientists, like most scientists, are a conservative lot. They try to use the minimum number of physiological mechanisms to explain all visual thresholds. Fifty years

ago, Hecht and his colleagues ascribed everything from dark adaptation to acuity to the properties of the photoreceptors (Hecht, 1937). For the last two decades, the trend has been to invoke the properties of neurones in the primary visual cortex to account for numerous psychophysical thresholds (Thomas, 1970). Many contemporary psychophysical models propose hypothetical spatial channels or mechanisms, analogous to the more familiar colour mechanisms that form the theoretical basis of colour vision models. Each spatial mechanism consists of a set of identical units that differ only in their spatial locations (Wilson, 1991); these units, modelled as spatial filters of a particular size, correspond roughly to the simple cells described originally by Hubel and Wiesel (1968). Multiple mechanisms, each based on a different sized unit (or filter), were used initially to explain the abundant data on contrast sensitivity, but more recently, these mechanisms have been called upon to explain hyperacuity thresholds (Klein and Levi, 1985; Nielsen *et al.*, 1985; Wilson, 1986; 1991). Because of the significant role of contrast in these models, there have been many measurements of the effect of contrast on the hyperacuities.

Why should contrast have anything to do with spatial position? The hypothetical link between these two dimensions comes from the response characteristics of single cortical units. For a target positioned within the receptive field of a simple cell, changes in target contrast, orientation and position can all produce equivalent changes in the response of the cell. Thus, these dimensions are necessarily confounded by the response of a single cell, but not, of course, by the human observer. Every position in the visual field is represented by many cortical units with overlapping receptive fields, so from the perspective of all the cells responding to the target, the overall pattern of cortical signals generated by a change in contrast will be quite different from that generated by a shift in position. Some subsequent stage of processing could easily differentiate between the two, but no subsequent processing, no matter how clever, can identify the signal if the signal is not detected. If these simple cortical units are the site of the noise limiting signal detection, then the minimum detectable change in contrast and the minimum detectable change in position are related in a straightforward manner (Nakayama and Silverman, 1985; see Shapley and Victor, 1986 for a similar discussion of ganglion units). For psychophysicists, there is no need to assume a particular physiological site; the key element in the psychophysical models is that contrast discrimination and position sensitivity are limited by the same source of noise whatever its physiological origin (Klein and Levi, 1985; Morgan and Aiba, 1985a; Regan and Beverley, 1985; Morgan, 1986; Wilson, 1986; Bowne, 1990).

The simplistic diagram in Fig. 17.5 shows one conceivable relationship between the contrast increment threshold and the position threshold. (This simple diagram owes much to the thoughtful analysis proposed by Morgan (Morgan and Aiba, 1985a; Morgan, 1986) in his extension of the Hartridge model. The errors in conception are, of course, my own. It is important to note that Morgan's approach did not include 'spatial filters'.) The curves at the bottom of the figure are the one-dimensional profiles (weighting functions) of the spatial filters that represent cortical units in the psychophysical models. The pictures above these putative filters represent the luminance profiles of bar stimuli at two contrast values, the uppermost row representing a contrast twice the value of the row below it. The left half of the figure shows the incremental change in the bar contrast that produces some criterion

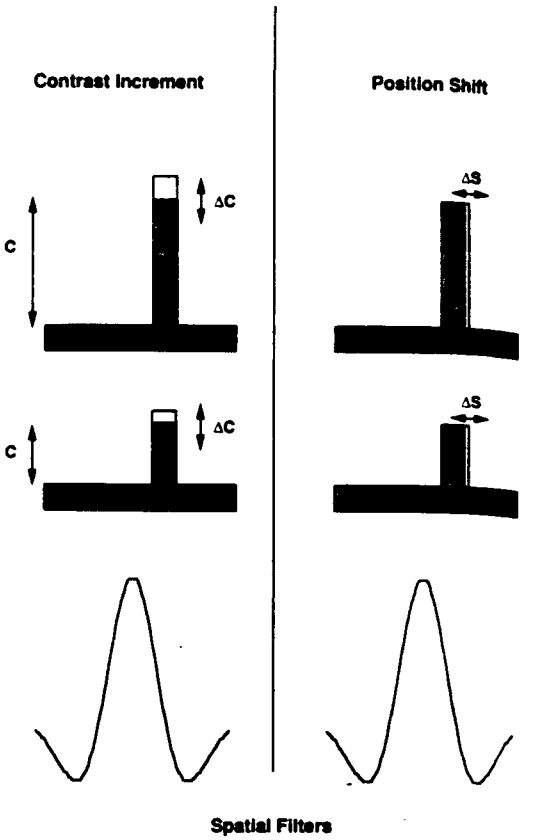


Fig. 17.5 *Diagram of relation between incremental changes in contrast and incremental shifts in position. Lower row shows one-dimensional profile of spatial filters, the psychophysical representation of simple cortical units. Middle row left shows contrast increment that produces detectable response change; middle row right shows spatial shift that produces equivalent response change. Upper row left shows contrast increment for twice the contrast for* $\Delta C/C = constant$ *(Weber's law); upper row right shows that position shift will be independent of contrast if Weber's law is true. See text for details.*

change in the filter response, while the right half of the figure shows the positional shift that produces the same change in the filter response. Suppose that the two responses have been equated at the contrast level in the middle row. What happens if the contrast is doubled as shown in the top row? If Weber's law were true for contrast discrimination ($\Delta C/C$ = a constant), then the contrast increment would double, but the position threshold would not change. This occurs because a shift of the same size moves a bar of twice the contrast, so the incremental change in response is doubled, matching the incremental change in contrast for the higher contrast level. Thus, if contrast and position increments are limited by the same source of noise, and contrast discrimination obeys Weber's law, position thresholds should be independent of contrast.

In fact, contrast discrimination does not obey Weber's law. The Weber fraction $\Delta C/C$ gradually decreases with increasing contrast indicating an improvement in the signal-to-noise ratio. It typically falls with a slope of -0.4 to -0.5 on log-log coordinates, a value found for many different types of targets, both stationary and moving (Legge and Foley, 1980; Legge, 1981; Legge and Kersten, 1983; Bowne, 1990). Shallower slopes are observed for briefly presented targets (Legge and Kersten, 1983), and steeper slopes are reported for high-frequency band-limited targets (Kulikowski, 1976; Kulikowski and Gorea, 1978; Wilson, 1991). Thus, following the same reasoning demonstrated in Fig. 17.5, positional thresholds should decline with increasing contrast, mimicking the decrease in the contrast Weber fraction.

Some of the hyperacuities do indeed improve with increasing contrast in a way quite consistent with the assumption that contrast and position thresholds are limited by a common source of noise. All reported measurements of Vernier acuity improve with contrast. Most studies find that Vernier thresholds fall with a slope -0.5 on log-log coordinates (Watt and Morgan, 1983; Krauskopf and Farell, 1990; Wehrhahn and Westheimer, 1990), although a few studies have reported steeper slopes (Wilson, 1986; Bradley and Skottun, 1987). A representative graph showing the effect of contrast on Vernier acuity is plotted in Fig. 17.6(a); the data points are from a recent study by Kiorpes and Movshon (1990). The graphs in Fig. 17.6 show that stereoacuity (Legge and Gu, 1989), curvature detection (Hess and Watt, 1990) and motion displacement (Mather, 1987) also improve with contrast with nearly the same -0.5 slope found in the Vernier studies. Stereoacuity has the additional requirement that the contrast in the two eyes be equal for best performance

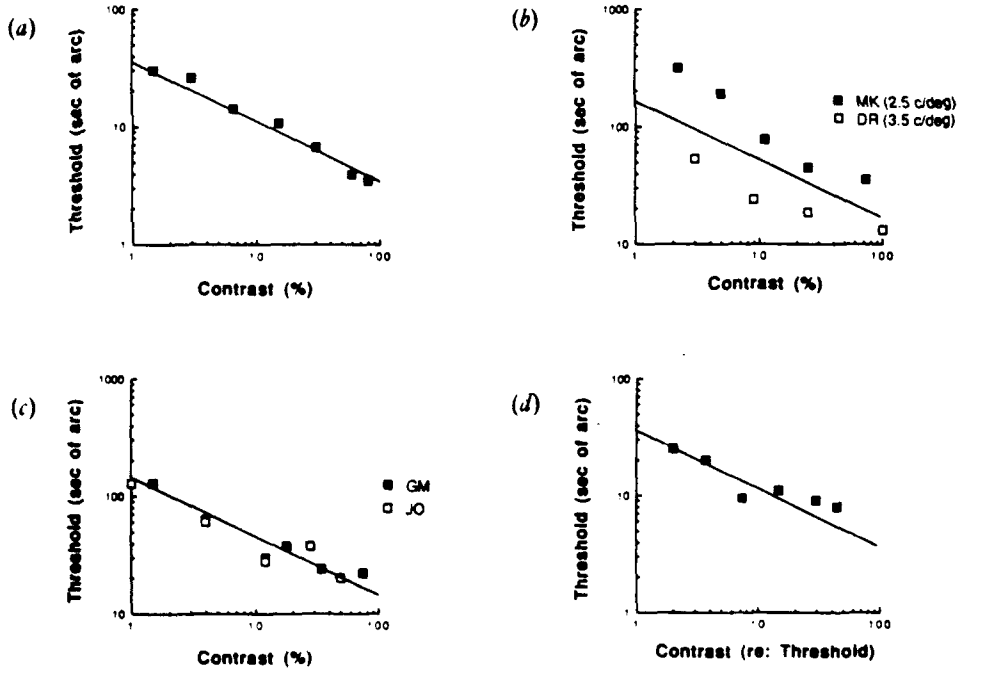


Fig. 17.6 *Four hyperacuity judgments that improve with contrast. Continuous lines have -0.5 slope. (a) Vernier acuity for sinusoidal grating (4.6 c deg^{-1}); data replotted from Kiorpes and Movshon, 1990. (b) Stereoacuity for sinusoidal grating; data replotted from Legge and Gu, 1989. (c) Motion displacement thresholds for step edge; data replotted from Mather, 1987. (d) Curvature thresholds for detecting deviation ('bump') from straightness in line presented in fovea. (Data replotted from Hess and Watt, 1990.)*

(Halpern and Blake, 1988; Heckmann and Schor, 1989; Legge and Gu, 1989). If a high contrast target in one eye is paired with a low contrast target in the other, the stereo threshold is higher than if both eyes have the same low contrast.

The relationship between contrast and position is probably not as simple as Fig. 17.5 suggests. For one thing, careful concurrent measurements of contrast discrimination and Vernier thresholds with identical target configurations indicate that the contrast dependence of these two measurements is never quite the same; $\Delta C/C$ always falls with slope close to -0.5, while the Vernier thresholds always show a slightly steeper decline with increasing contrast (Hu *et al.*, 1990). For another, the contrast dependence is not identical for all hyperacuities. Westheimer and Pettet (1990) compared the effect of contrast on Vernier and stereoacuity thresholds using the same target. As Fig. 17.7(a) shows, both the Vernier and stereo acuity thresholds decrease with increasing contrast, but the stereo function is shifted laterally with respect to the Vernier function as though the effective contrast for stereo were half the contrast for the Vernier configuration.

Finally, not all positional thresholds show the predicted improvement with contrast. The most puzzling exceptions are spatial interval judgments. In a spatial interval task, the observer judges whether the test separation is less or greater than some baseline value. At baseline separations greater than about 3′ in the fovea, spatial interval judgments are not affected by target contrast once contrast exceeds 2–3 times threshold (Levi and Klein, personal communication). Morgan and Regan (1987) measured both Vernier and spatial interval thresholds using thin bars with a Gaussian luminance profile. As Fig. 17.7(b) demonstrates, the Vernier thresholds improve with contrast, while the spatial interval judgments for a base separation of 5′ are nearly independent of contrast. A related judgment, spatial frequency discrimination, in which the observer judges the separation between the bars of a sinusoidal grating, also shows no improvement with contrast at levels above 2–3% contrast (Regan *et al.*, 1982; Thomas, 1983; Skottun *et al.*, 1987; Bowne, 1990). Many positional thresholds measured with grating stimuli show this same contrast independence. The thresholds for motion displacement (Nakayama and Silverman, 1985) and grating orientation (Bowne, 1990) plotted in Fig. 17.7(c) and (d) change almost imperceptibly with increasing contrast

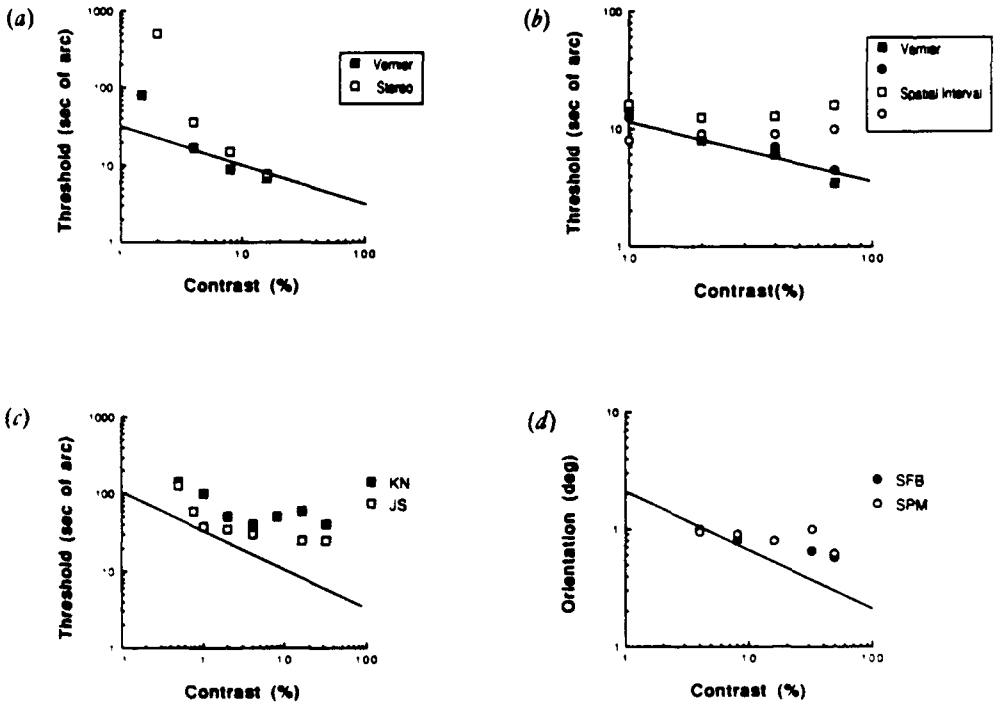


Fig. 17.7 *The effect of contrast is not the same on all positional thresholds. Continuous lines have -0.5 slope. (a) Vernier acuity and stereoacuity thresholds for identical block targets; data replotted from Westheimer and Pettet, 1990. (b) Vernier acuity and spatial interval thresholds for thin lines with Gaussian profile; data replotted from Morgan and Regan, 1987. (c) Motion displacement thresholds for sinusoidal grating (2 c deg^{-1}); data replotted from Nakayama and Silverman, 1985. (d) Orientation thresholds for sinusoidal grating (4 c deg^{-1}). (Data replotted from Bowne, 1990.)*

beyond 2%. Similar results on orientation have been reported by Regan and Beverley (1985) and Skottun *et al.* (1987).

Wavelength

There have been relatively few studies of the effect of wavelength on hyperacuity. Baker (1949) measured Vernier acuity for a dark line presented against various coloured backgrounds and found slightly poorer performance with a blue background (peak wavelength (490 nm)) and slightly enhanced performance with a red (690 nm) or a yellow background (575 nm) when compared to a white background of the matched brightness. She attributed the decrement in blue light to the fact that her observers were not in focus for blue light in her apparatus, and demonstrated that the addition of appropriate lens power greatly improved the thresholds for a blue background to values equal or better than the other wavelength conditions. Perhaps some small amount of defocus also accounts for the small decrement in Vernier acuity in blue light noted by Foley-Fisher (1973).

Morgan and Aiba (1985a) created a Vernier target by manipulating the relative luminance of two adjacent unresolved bars. The luminance centroid in the upper half of the two bars was moved with respect to the centroid in the lower half of the bars, so that the target appeared to be spatially offset. They then compared the detectable 'offset' for a pair of unresolved white light bars to a pair of unresolved coloured bars (red and green). Colour made no essential difference to the task, since the apparent offset depended only on the luminance ratio of the adjacent bars once the equivalent luminance of the red and green bars had been taken into account. Morgan and Aiba (1985b) also looked at Vernier acuity with equiluminant bars (bars with only colour contrast and no contrast in luminance). The thresholds for the equiluminous condition were higher than those for a comparable luminance contrast, a result they conjectured was caused by the shift from bandpass filtering to low-pass filtering in the opponent-colour units responding to the equiluminous target – a neural change that is similar to blurring the physical stimulus.

Exposure Duration

Does hyperacuity get better if you have more time to look at the target? It might seem that these highly precise judgments would require an extended period of target scrutiny. Oddly enough, increasing exposure duration does not greatly improve Vernier acuity for high luminance line targets (Westheimer and McKee, 1977a). For small or dim targets, increasing duration improves Vernier thresholds largely because of the temporal summation of light. Hadani *et al.* (1984) measured three-dot Vernier alignment as a function of exposure duration. In a three-dot alignment task, the observer judges whether the central dot is displaced to the right or left of the implicit line defined by the outer two reference dots. For a constant luminance target, they found a modest improvement with duration, the thresholds falling with a slope of −0.25, reminiscent of the improvement with increasing luminance noted above for some types of hyperacuity target. Indeed, Hadani *et al.* argued that the decrease in the thresholds was entirely due to temporal energy summation, an argument supported by their demonstration that the threshold for a constant energy target (luminance × time = constant) did not change with increasing

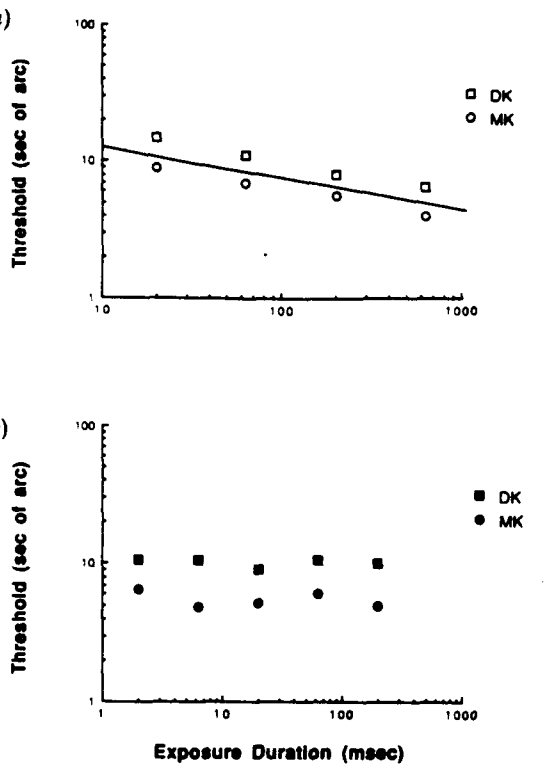


Fig. 17.8 *Part of the improvement in hyperacuity thresholds with exposure duration can be attributed to the temporal summation of light. (a) Vernier acuity thresholds for a constant luminance targets. Continuous line has slope of −0.25, equal to the slope for luminance dependence of hyperacuity. (b) Vernier acuity thresholds for constant energy (luminance × time = constant) targets. (Replotted from Hadani* et al., *1984.)*

exposure duration. Their results are replotted in Fig. 17.8.

Nevertheless, energy summation is not the whole story. Subtle changes in hyperacuity thresholds are noted even when duration is increased from 150–1000 ms and it is difficult to attribute changes over this time frame to temporal energy summation, particularly because more improvement is found in foveal thresholds than in peripheral thresholds (Yap *et al.*, 1987a). Burbeck (1986) found that spatial interval thresholds for band-limited high frequency 'bars' improved more with increasing duration than broad-band targets (ordinary bars). Echoing an earlier suggestion by Morgan *et al.* (1983), Burbeck concluded that the signal-to-noise ratio for small scale (high frequency) mechanisms takes more time to reach its steady-state value. This conclusion is difficult to reconcile with the observation that resolution acuity is not affected significantly by exposure duration (Stigmar, 1971b; Watt, 1987).

Exposure duration does not have a uniform effect on all hyperacuities. Stereoacuity seems to require an especially long stimulus duration (Ogle and Weil, 1958). Keesey (1960) measured Vernier thresholds as a function of exposure duration, and one of her observer (G.K.S.) was a subject for similar measurements on stereoacuity (Shortess and Krauskopf, 1961). These data have been replotted in Fig. 17.9(a). The Vernier thresholds appear to reach an asymptotic value at a shorter duration than the stereo thresholds which are still improving at a duration of 1 s. Measurements by Stigmar (1971b) of Vernier and stereo thresholds show a similar pattern. Foley and Tyler (1976) found parallel improvement in stereo and Vernier thresholds with increasing duration, but the stereo thresholds were systematically higher. Westheimer and Pettet (1990) used much brighter targets to measure stereoacuity and Vernier acuity concurrently; representative data from their study are replotted in the lower graph of Fig. 17.9. Their Vernier thresholds for these bright targets are almost independent of exposure duration, while the stereo thresholds for the identical target configuration are still improving at 500 ms.

Why does stereoacuity require more time than Vernier acuity? Since each of the studies cited above used a different target configuration and a different luminance level, it is doubtful that good stereopsis simply demands more light than Vernier acuity. The best stereoacuity is found for targets presented in the fixation plane, so perhaps observers must gently modulate their convergence during the target presentation to find the optimum position, a search that should certainly take time. The most likely explanation is that the neural processing time associated with stereopsis is substantially longer than that needed for the two-dimensional hyperacuities.

Watt (1987) made detailed measurements of several hyperacuities as a function of exposure duration. He followed each target presentation with a masking stimulus to obscure any persisting signal from the retinal image. This masking paradigm appears to disrupt ordinary hyperacuity processing. All of Watt's hyperacuity thresholds (curvature, line width, orientation, stereoacuity) improved dramatically with increasing duration even though Watt had kept target energy constant for durations shorter than 100 ms. Most of the thresholds fell 1–1.5 log units as duration increased from 10–1000 ms, but the stereoacuity thresholds declined by more than 2 log units. Concurrent measurements of resolution acuity showed very slight improvement for the same conditions. Stigmar (1971b) also had noted that stereo and Vernier thresholds took more time to reach asymptotic performance than comparable measures of resolution acuity.

The hyperacuity thresholds for bright, high contrast targets viewed in isolation are roughly independent of duration, stereoacuity being a notable exception. Watt's results suggest that the processing of the shape informa-

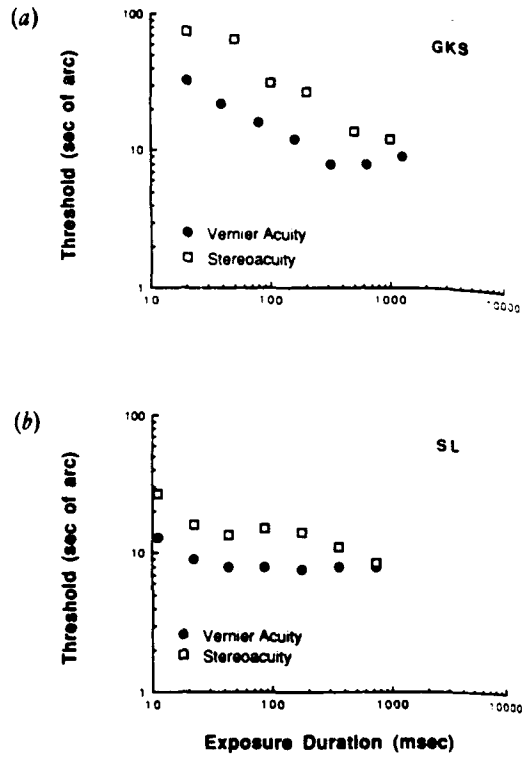


Fig. 17.9 *Exposure duration has differential effects on stereoacuity and Vernier acuity. (a) Data replotted from two different studies (Keesey, 1960; Shortess and Krauskopf, 1961) that used the same subject for stereo and Vernier measurements with line targets. (b) Stereo and Vernier acuity measurement with an identical block targets. Luminance was roughly three times the level used for data in (a). (Data replotted from Westheimer and Pettet, 1990.)*

tion that mediates these thresholds can be significantly disrupted by temporal events that follow the target in rapid succession. The results of Westheimer and Hauske (1975) also support the notion that hyperacuity requires some period of unimpeded processing time for best performance.

Eccentricity

Strictly speaking, eccentricity is a physiological, not a physical, constraint. As a practical matter, one often has control over where in the visual field that the target will be displayed so it is useful to consider the effects of eccentricity on performance. Predictably, hyperacuity thresholds increase with increasing eccentricity, reflecting well-known differences between the processing capabilities of the human fovea and periphery. The interesting fact is that hyperacuity thresholds rise more rapidly with eccentricity than do resolution acuity thresholds (Westheimer, 1982; Fendick and Westheimer, 1983; McKee and Nakayama, 1984; Fahle and Schmid, 1988).

In 1985, Levi *et al.* observed that all the published data on hyperacuity showed a similar decline with eccentricity. When the peripheral thresholds were scaled against the foveal value, the hyperacuity eccentricity functions, plotted on linear axes, intercepted the x-axis (the eccentricity axis) at values ranging from -0.5 to -1.0, while the eccentricity functions for resolution typically intercepted the x-axis at values ranging between -2.0 and -3.0. Recently McKee *et al.* (1990) measured spatial interval, stereoacuity and motion displacement thresholds in the same subjects with identical target configurations. They found a nearly identical rise with eccentricity in all three hyperacuity thresholds for measurements made in the perifoveal region; Fig. 17.10 (top) shows representative data replotted from their study. Although generally hyperacuity thresholds appear to be limited by some common factor that changes systematically with eccentricity, Foster *et al.* (1989) did find small systematic differences (about a factor or two) between spatial interval and motion displacement thresholds measured beyond 10° eccentricity. Another curious observation is that, at the same eccentric locus, judgments of spatial position made along a radial axis extending from the fovea are poorer than judgments made along the tangent to a circle centred on the fovea (Yap *et al.*, 1987; Klein and Levi, 1987).

Stereoacuity has an additional important constraint. The stereoacuity thresholds plotted in Fig. 17.10(a) were obtained for targets presented on the horopter (in or very near the fixation plane). Stereoacuity thresholds rise rapidly as the target is moved off the horopter (Ogle, 1953; Blakemore, 1970; Westheimer and McKee, 1978; Badcock and Schor, 1985). As the distance from the horopter is increased, the standing disparity of the target lines obviously increases with the result that the eccentricity of the stereo half-images (the images in the two eyes that are fused by the cortex resulting in a depth percept) must also increase. Schor and Badcock (1985) thought that perhaps the eccentricity of the half-images could explain the decrement in off-horopter stereoacuity. To test this idea, they compared off-horopter stereoacuity to vernier acuity at an equivalent eccentricity and demonstrated that the loss in stereoacuity is not due to eccentricity of the half images.

Disparity is a distance, the dichoptic distance between the locations of the stereo half-images in the two eyes. McKee *et al.* (1990) measured spatial interval and stereo thresholds as a function of the distance or separation between the monocularly-presented lines of the spatial interval targets and of the disparity of the dichoptically-

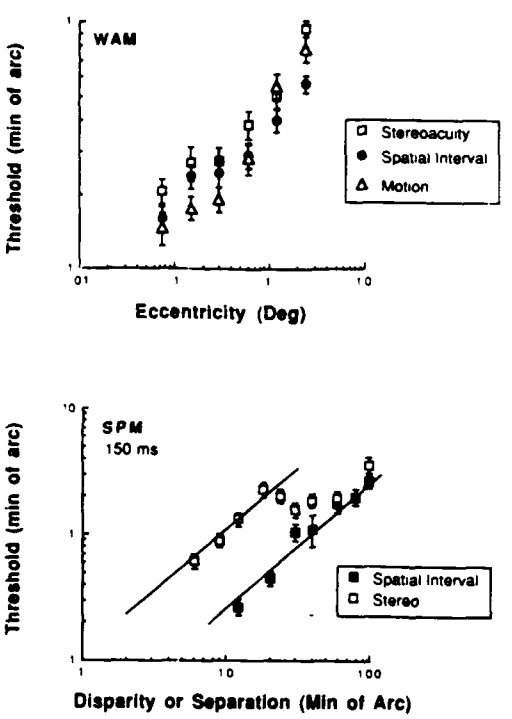


Fig. 17.10 *(a) All hyperacuities show a similar decline with eccentricity in the perifoveal region. Three measurements of hyperacuity (stereoacuity, spatial interval amd motion displacement thresholds) for identical targets measured as a function of eccentricity. Data replotted from McKee* et al., *1990b. (b) Off-horoptre stereo judgments are substantially less precise than comparable judgments for spatial interval. Spatial interval thresholds measured as function of separation between target lines. Disparity thresholds measured as function of standing (or pedestal) disparity. (Data replotted from McKee* et al., *1990a.)*

presented lines of the stereo target. As is apparent from Fig. 17.10(b), the stereo thresholds are substantially higher than the spatial interval thresholds for disparities ranging up to about 40′. At larger disparities, the half-images of the stereo target are quite diplopic and the associated depth is weak. Interestingly, at these same large disparities, the thresholds for spatial interval and for stereo are nearly identical – 'dichoptic' spatial interval judgments, made when the target lines defining the spatial interval are presented to different eyes, are as precise as spatial interval judgments made when both target lines are presented to the same eye.

Blur

The best hyperacuity is found for targets in sharp focus, but again the degradation produced by defocus depends on the task. A few studies have measured the influence of spectacle blur on hyperacuity, but most investigators have preferred to use ground-glass blur or Gaussian blur because the resulting images are more amenable to psychophysical analysis. For example, Westheimer and McKee (1980) found that stereoacuity was more degraded than resolution by spectacle blur, but the phantom targets introduced by spurious resolution with dioptric defocus could degrade stereoacuity for reasons that have little to do with the 'fuzziness' of the target. Westheimer (1979) found that motion displacement thresholds were less affected than resolution for small amounts of dioptric defocus; McKee and Nakayama (1984) also found that dioptric defocus had little effect on motion displacement thresholds.

Even in the simpler regime represented by ground-glass or Gaussian blur, the effects of blur are far from simple. Consider what blur does to the targets shown in Fig. 17.1. Blur will decrease target contrast, so if the measured hyperacuity is sensitive to contrast then the change in contrast alone will degrade performance. For targets formed of adjacent lines or points, blur can 'merge' the local features obscuring relative position. In their careful study of the effect of ground glass blur on Vernier acuity, Williams *et al.* (1984) identified configurations (abutting lines or two points separated by a small gap) where blur produced a dramatic increase in thresholds, but they were also able to identify targets (two points with a large gap) where the effect of blur had little effect on alignment thresholds. Stigmar (1971a) noted a similar dependence on target configuration when he measured the effect of ground glass blur on stereo- and Vernier acuity. Unfortunately, the best hyperacuity is generally obtained when targets are fairly close together, so the overriding fact is that blur degrades thresholds.

Some studies have sought to analyse the effect of target spread *per se*. Toet *et al.* (1987) measured three 'blob' Vernier acuity and bisection acuity at contrast threshold, defined as that contrast where 80% of the targets were seen. The separation between the features was increased as the spread of the Gaussian blurred targets increased. Under these conditions, thresholds were a linear function of increasing blur, for blobs with a spread greater than 1.5′. This approach necessarily confounds separation with target blur, but Toet *et al.* were able to analyse the distinct contributions of each factor, concluding that the relative importance of target blur and target separation depended on the task.

An even more elaborate analysis of the effect of blur was performed by Levi and Klein (1990b) with a different objective. Some neural operations degrade ideal performance in a way that might be modelled as a kind of internal or 'intrinsic' blur. Levi and Klein measured this intrinsic blur by increasing the physical blur until two-point resolution thresholds were affected by the stimulus blur. They then examined the effect of blur on other tasks (line detection, spatial interval discrimination). In the context of the effect of blur on hyperacuity, several of their conclusions are noteworthy. For a wide range of separations and eccentricities, spatial interval thresholds rise when stimulus blur (or spread) exceeds about one third to one half the distance between the blurred features. The measured intrinsic blur varies with eccentricity in a way consistent with the cone spacing in the central 10° of the visual field, as does resolution acuity. Thus, intrinsic blur does not explain why hyperacuity thresholds fall off more rapidly with eccentricity than does resolution.

Target Length

Traditionally, foveal hyperacuity thresholds were explained by some process that averaged the 'local signs' (locations) of the cones stimulated by the extended contours of the lines or edges forming the target (Hering, 1899; Weymouth *et al.*, 1923). It is now abundantly clear that no such process is needed to explain hyperacuity thresholds (Westheimer, 1976; Geisler, 1984). An ideal detector could localize a point source to an arbitrary precision provided that there was enough light in the image. Certainly Geisler's calculations indicate that an ideal observer limited only by the optics of the eye and the properties of the photoreceptors could localize a point source to a much higher precision than that reached by real observers.

Nevertheless, human thresholds should improve as the length of the target lines increases, if only because there are more chances for the underlying neural units to detect

a small change in position. Unfortunately, this prediction is incorrect. In the fovea, hyperacuity thresholds can be as low for properly-positioned points as they are for lines of any length chosen (Ludvigh, 1953; Sullivan *et al.*, 1972; Westheimer and McKee, 1977a). This result means that the human visual system is making inefficient use of the spatial information because there is necessarily more information in lines than in points. Indeed, some hyperacuity thresholds (Vernier acuity for abutting line targets, curvature) do improve with increasing line length so, in some sense, the relative efficiency of human performance increases for these tasks (Andrews *et al.*, 1973). Why *any* hyperacuity thresholds should be as low for point targets as for lines remains a mystery.

Outside of the fovea, line length matters. Andrews and Miller (1978) and Levi and Klein (1986) found a significant improvement in bisection thresholds with increasing length even at small perifoveal distances (1.3° eccentricity and 2.5° eccentricity). A common strategy for hyperacuity measurements in the periphery is to scale the line length according to the hyperacuity eccentricity function from Levi *et al.* (1985):

$$\text{Length} = x(1 + E/0.77) \qquad (17.1)$$

where E is the eccentricity given in degrees and x is the length of the foveal stimulus. Whether this strategy is appropriate for all hyperacuities is unknown, so perhaps preliminary measurements on the effect of line length should be a first step in experimental studies of peripheral processing.

Concluding Comments

What do hyperacuity thresholds measure? Although these highly precise thresholds are limited by light and contrast in fairly predictable ways, human observers can easily discriminate between threshold changes in contrast (or light), and threshold shifts in position (McKee *et al.*, 1990). Therefore, these local spatial changes must be given neural 'labels' that are different from the 'labels' applied to contrast or luminance changes (Westheimer, 1979). What aspect of pattern or object perception is encoded by the labelled changes that we call hyperacuity? Since the experimenter manipulates the location of a some feature in measuring these thresholds, it might seem reasonable to think of hyperacuity as a measure of the ability to identify visual location, but I am convinced that hyperacuity is a measure of shape, not location. (This idea was suggested to me by Sam Bowne, although it is also implicit in the work of Geisler (1984), Levi *et al.* (1988), Burbeck and Yap (1990) and Levi and Klein (1990b).) This statement is not intended as a semantic quibble. I recognize that one cannot change shape without manipulating the relative location of the defining contours, but I contend that we do not know the position of any given feature to the precision associated with hyperacuity thresholds.

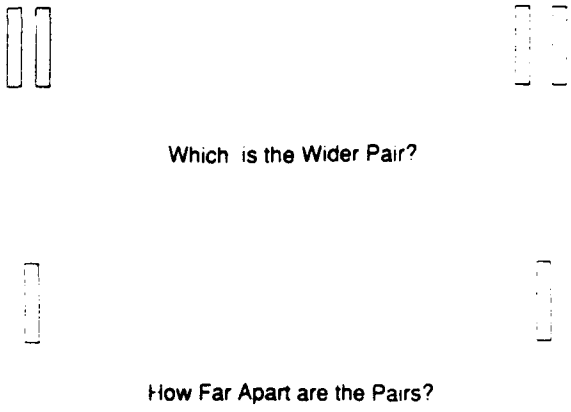


Fig. 17.11 *The thresholds for judging the relative width of the upper pairs of lines are much smaller than the thresholds for judging the distance between the pairs. Local shape judgments are more precise than judgments of location.*

Consider the targets in Fig. 17.11. If I ask you to discriminate the relative widths of the two pairs of lines at the top of the figure, you will be able to judge which is wider with hyperacuity precision, even if the targets are presented too briefly to allow you to move your eyes. On the other hand, if I ask you to judge the distance separating the pairs, your judgments will be much less precise (McKee *et al.* 1990b). You have much better information about the local shape of the contours, than you have about their location. In this sense, hyperacuity is a primitive measure of shape perception. This chapter has described the influence of luminance, contrast, exposure duration and numerous other physical variables on hyperacuity *per se*, but, in a larger context, the same conditions that produce optimum performance in hyperacuity judgments will also ensure that shapes and text can be easily interpreted.

Acknowledgements

I am deeply indebted to Dr Samuel F. Bowne for his many informal tutorials on contemporary hyperacuity models and for his imaginative and thoughtful discussions of many of the issues described in this chapter. I also want to thank Dr Leslie Welch for many years of bright, lively, thought-provoking discussions on all aspects of hyperacuity. I am grateful to the many colleagues who supplied reprints and preprints of their own work. Work on this chapter was supported by NEI Grant RO1-EY06644, AFOSR-89-0035 and NEI Core Grant 5P-30-EY01186.

References

Andrews, D. P. and Miller, D. T. (1978). Acuity for spatial separation as a function of stimulus size. *Vision Res.*, **18**, 615–619.

Andrews, D. P., Butcher, A. K. and Buckley, B. R. (1973). Acuities for spatial arrangement in line figures: Human and ideal observers compared. *Vis. Res.*, **13**, 599–620.

Badcock, D. R. and Schor, C. M. (1985). Depth-increment detection function for individual spatial channels. *J. Opt. Soc. Am. A*, **2**, 1211–1215.

Baker, K. E. (1949). Some variables influencing Vernier acuity: I. Illumination and exposure time. II. Wave-length of illumination. *J. Opt. Soc. Am.*, **39**, 567–576.

Berry, R. N. (1948). Quantitative relations among vernier, real depth, and stereoscopic depth acuities. *J. Exp. Psychol.*, **38**, 708–721.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *J. Physiol.*, **211**, 599–622.

Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Res.*, **30**, 449–462.

Bradley, A. and Skottun, B. C. (1987). Effects of contrast and spatial frequency on vernier acuity. *Vision Res.*, **27**, 1817–1824.

Burbeck, C. A. (1986). Exposure-duration effects in localization judgments. *J. Opt. Soc. Am. A*, **3**, 1983–1988.

Burbeck, C. A. and Yap, Y. L. (1990). Two mechanisms for localization? Evidence for separation-dependent and separation-independent processing of position information. *Vision Res.*, **30**, 739–750.

Fahle, M. and Schmid, M. (1988). Naso-temporal asymmetry of visual perception and of the visual cortex. *Vision Res.*, **23**, 293–300.

Fendick, M. and Westheimer, G. (1983). Effects of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vision Res.*, **23**, 145–150.

Foley, J. M. and Tyler, C. W. (1976). Effect of stimulus duration on stereo and vernier displacement thresholds. *Percept. Psychophys.*, **20**, 125–128.

Foley-Fisher, J. A. (1973). The effect of target line length on vernier acuity in white and blue light. *Vision Res.*, **13**, 1447–1454.

Foster, D. H., Gravano, S. and Tomoszek, A. (1989). Acuity for fine-grain motion and for two-dot spacing as a function of retinal eccentricity: differences in specialization of the central and peripheral retina. *Vision Res.*, **29**, 1017–1031.

Geisler, W. (1984). Physical limits of acuity and hyperacuity. *J. Opt. Soc. Am. A*, **1**, 775–782.

Geisler, W. (1985). Ideal discriminators in spatial vision: two-point stimuli. *J. Opt. Soc. Am. A*, **2**, 1483–1497.

Geisler, W. (1989). Sequential ideal-observer of visual discriminations. *Psychol. Rev.*, **96**, 276–314.

Geisler, W. and Davia, R. D. (1985). Ideal discriminators in spatial vision: two-point stimuli. *J. Opt. Soc. Am. A*, **2**, 1483–1497.

Hadani, I., Meiri, A. Z. and Guir, M. (1984). The effects of exposure duration and luminance on the 3-dot hyperacuity task. *Vision Res.*, **24**, 871–874.

Halpern, D. L. and Blake, R. R. (1988). How contrast affects stereoacuity. *Perception*, **17**, 483–495.

Hecht, S. (1937). Rods, cones, and the chemical basis of vision. *Physiol. Rev.*, **17**, 255–281.

Heckmann, T. and Schor, C. M. (1989). Is edge information for stereoacuity spatially channeled? *Vision Res.*, **29**, 593–607.

Hering, E. (1889). Über die Grenzen der Sehschärfe. *Ber. Math. Phys. Classe Königl. Sachs. Ges. Wissenschaft (Leipzig)*, **16–24**.

Hess, R. F. and Watt, R. J. (1990). Regional distribution of the mechanisms that underlie spatial localization. *Vision Res.*, **30**, 1021–1031.

Hu, Q., Klein, S. A. and Carney, T. (1990). Vernier acuity and contrast discrimination of sinusoidal gratings. *Invest. Ophthalmol. Vis. Sci.*, **31**, 327.

Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.*, **195**, 215–243.

Keesey, U. T. (1960). Effects of involuntary eye movements on visual acuity. *J. Opt. Soc. Am.*, **50**, 769–774.

Kiorpes, L. and Movshon, J. A. (1990). The effect of contrast on vernier acuity. *Invest. Ophthalmol. Vis. Sci.*, **31**, 279.

Klein, S. A. and Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation. *J. Opt. Soc. Am. A*, **2**, 1170–1190.

Klein, S. A. and Levi, D. M. (1987). Position sense of of the peripheral retina. *J. Opt. Soc. Am. A*, **4**, 1543–1553.

Krauskopf, J. and Farell, B. (1990). Vernier acuity: effects of chromatic content, blur and contrast. *Vision Res.* (in press).

Kulikowski, J. J. (1976). Effective contrast constancy and linearity of contrast sensation. *Vision Res.*, **16**, 1419–1431.

Kulikowski, J. J. and Gorea, A. (1978). Complete adaptation to patterned stimuli: A necessary and sufficient conditions for Weber's law for contrast. *Vision Res.*, **18**, 1223–1227.

Legge, G. E. (1981). A power law for contrast discrimination. *Vis. Res.*, **21**, 457–467.

Legge, G. E. and Campbell, F. W. (1981). Displacement detection in human vision. *Vision Res.*, **21**, 205–213.

Legge, G. E. and Foley, J. M. (1980). Contrast masking in human vision. *J. Opt. Soc. Am.*, **70**, 1458–1471.

Legge, G. E. and Gu, Y. (1989). Stereopsis and contrast. *Vis. Res.*, **29**, 989–1004.

Legge, G. E. and Kersten, D. (1983). Light and dark bars: Contrast discrimination. *Vision Res.*, **23**, 473–483.

Levi, D. M. and Klein, S. A. (1990a). The role of separation and eccentricity in encoding position. *Vision Res.*, **30**, 557–585.

Levi, D. M. and Klein, S. A. (1986). Sampling in spatial vision. *Nature*, **320**, 360–362.

Levi, D. M. and Klein, S. A. (1990b). Equivalent intrinsic blur in spatial vision. *Vision Res.*, (in press).

Levi, D. M., Klein, S. A. and Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *Vision. Res.*, **25**, 963–977.

Levi, D. M., Klein, S. A. and Yap, Y. L. (1988). 'Weber's law' for position: unconfounding the role of separation and eccentricity. *Vision Res.*, **28**, 597–603.

Ludvigh, E. (1953). Direction sense of the eye. *Am. J. Ophthalm.*, **36**, 139–142.

McKee, S. P. and Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Res.*, **24**, 25–32.

McKee, S. P., Levi, D. M. and Bowne, S. F. (1990a). The imprecision of stereopsis. *Vision Res.*, **30**, 1763–1780.

McKee, S. P., Welch, L., Taylor, D. G. and Bowne, S. F. (1990b). Finding the common bond: Stereoacuity and the other hyperacuities. *Vision Res.*, **30**, 879–891.

Mather, G. (1987). The dependence of edge displacement thresholds on edge blur, contrast and displacement distance. *Vision Res.*, **27**, 1631–1637.

Morgan, M. J. (1986). The detection of spatial discontinuities and interactions between contrast and spatial contiguity. *Spatial Vision*, **1**, 291–303.

Morgan, M. J. and Aiba, T. S. (1985a). Vernier acuity predicted from changes in the light distribution of the retinal image. *Spatial Vision*, **1**, 151–161.

Morgan, M. J. and Aiba, T. S. (1985b). Positional acuity with sinusoidal matic stimuli. *Vision Res.*, **25**, 689–695.

Morgan, M. J. and Regan, D. (1987). Opponent model for line interval discrimination: interval and vernier performance compared. *Vision Res.*, **27**, 107–118.

Morgan, M. J. and Watt, R. J. (1982). The modulation transfer function of a display oscilloscope: measurements and comments. *Vision Res.*, **22**, 1083–1085.

Morgan, M. J., Watt, R. J. and McKee, S. P. (1983). Exposure duration affects the sensitivity of vernier acuity to target motion. *Vision Res.*, **23**, 541–546.

Nakayama, K. and Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *J. Opt. Soc. Am. A*, **2**, 267–274.

Nakayama, K. and Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Res.*, **21**, 427–433.

Nielsen, K. R. K., Watson, A. B. and Ahumada, A. J. (1985). Application of a computable model of human spatial vision to phase discrimination. *J. Opt. Soc. Am. A*, **2**, 1600–1606.

Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images. *J. Opt. Soc. Am.*, **43**, 906–913.

Ogle, K. N. and Weil, M. P. (1958). Stereoscopic vision and the duration of the stimulus. *Arch. Ophthalmol.*, **48**, 4–12.

Regan, D. and Beverley, K. I. (1985). Postadaptation orientation discrimination. *J. Opt. Soc. Am. A*, **2**, 147–155.

Regan, D., Bartol, S., Murray, T. J. and Beverley, K. I. (1982). Spatial frequency discrimination in normal vision and in patients with multiple sclerosis. *Brain*, **105**, 735–754.

Riggs, L. A. (1965). Visual acuity. In *Vision and Visual Perception*. ed. Graham, C. H. pp. 321–349. NY: John Wiley & Sons.

Schor, C. and Badcock, D. (1985). A comparison of stereo and vernier acuity within spatial channels as a function of distance from fixation. *Vision Res.*, **25**, 1113–1119.

Shapley, R. and Victor, J. (1986). Hyperacuity in cat ganglion cells. *Science*, **231**, 999–1002.

Shortess, G. K. and Krauskopf, J. (1961). Role of involuntary eye movements in stereoscopic acuity. *J. Opt. Soc. Am.*, **51**, 555–559.

Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I. and Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: A comparison of single cells and behavior. *J. Neurophysiol.*, **57**, 773–786.

Sperling, G. (1971). The description and luminous calibration of cathode ray oscilloscope visual displays. *Behav. Res. Methods Instrumentation*, **3**, 147–151.

Stigmar, G. (1971a). Blurred visual stimuli. II. The effect of blurred visual stimuli on Vernier and stereo acuity. *Acta Ophthalmol.*, **49**, 364–379.

Stigmar, G. (1971b). Visual acuity with brief stimuli. *Acta Ophthalmol.*, **49**, 633–648.

Sullivan, G. D., Oatley, K. and Sutherland, N. S. (1972). Vernier acuity as affected by target length and separation. *Percept. Psychophys.*, **12**, 438–444.

Thomas, J. P. (1970). Model of the function of receptive fields in human vsion. *Psychol. Rev.*, **77**, 121–134.

Thomas, J. P. (1983). Underlying psychometric function for detecting gratings and identifying spatial frequency. *J. Opt. Soc. Am.*, **73**, 751–758.

Toet, A. van Eekhout, M. P., Simons, H. L. J. J. and Koenderink, J. J. (1987). Scale invariant features of differential spatial displacement discrimination. *Vision Res.*, **27**, 441–451.

Watt, R. J. (1984). Towards a general theory of the visual acuities for shape and spatial arrangement. *Vision Res.*, **24**, 1377–1386.

Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *J. Opt. Soc. Am. A*, **4**, 2007–2021.

Watt, R. J. and Morgan, M. J. (1983). The recognition and representation of edge blur: evidence for spatial primitives in human vision. *Vision Res.*, **23**, 1465–1477.

Wehrhahn, C. and Westheimer, G. (1990). How vernier acuity depends on contrast. *Exp. Brain Res.*, **80**, 618–620.

Westheimer, G. (1975). Visual acuity and hyperacuity. *Invest. Ophthalmol. Vis. Sci.*, **14**, 570–571.

Westheimer, G. (1976). Diffraction theory and visual hyperacuity. *Am. J. Opt. Physiol. Opt.*, **53**, 362–364.

Westheimer, G. (1979). The spatial sense of the eye. *Invest. Ophthalmol. Vis. Sci.*, **18**, 893–912.

Westheimer, G. (1982). The spatial grain of the perifoveal visual field. *Vision Res.*, **22**, 157–162.

Westheimer, G. (1985). The oscilloscopic view: retinal illuminance and contrast of point and line targets. *Vision Res.*, **25**, 1097–1103.

Westheimer, G. and Hauske, G. (1975). Temporal and spatial interference with Vernier acuity. *Vision Res.*, **15**, 1137–1141.

Westheimer, G. and McKee, S. P. (1977a). Integration regions for visual hyperacuity. *Vision Res.*, **17**, 89–93.

Westheimer, G. and McKee, S. P. (1977b). Spatial configurations for visual hyperacuity. *Vision Res.*, **17**, 941–947.

Westheimer, G. and McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *J. Opt. Soc. Am.*, **68**, 450–455.

Westheimer, G. and McKee, S. P. (1979). What prior uniocular processing is necessary for stereopsis? *Invest. Ophthalmol. Vis. Sci.*, **18**, 614–621.

Westheimer, G. and McKee, S. P. (1980). Stereoscopic acuity with defocused and spatially filtered retinal images. *J. Opt. Soc. Am.*, **70**, 772–778.

Westheimer, G. and Pettet (1990). Contrast and exposure duration differentially affect vernier and stereoscopic acuity. *Proc. R. Soc. Lond. B.* **241**, 42–46.

Weymouth, F. W., Andersen, E. E. and Averill, H. L. (1923). Retinal mean local sign: a new view of the relation of the retinal mosaic to visual perception. *Am. J. Physiol.*, **63**, 410–411.

Williams, R. A., Enoch, J. M. and Essock, E. A. (1984). The resistance of selected hyperacuity configurations to retinal image degradation. *Invest. Ophthalmol. Vis. Sci.*, **25**, 389–399.

Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Res.*, **26**, 453–470.

Wilson, H. R. (1991). Psychophysical models of spatial vision and hyperacuity. In *Vision and Visual Dysfunction*. Vol. 10. *Spatial Vision*. ed. Regan, D. London: Macmillan.

Yap, Y. L., Levi, D. M. and Klein, S. A. (1987a). Peripheral hyperacuity: three-dot bisection scales to a single factor from 0 to 10 degrees. *J. Opt. Soc. Am. A*, **4**, 1554–1561.

Yap, Y. L., Levi, D. M. and Klein, S. A. (1987b). Peripheral hyperacuity: isoeccentric bisection is better than radial bisection. *J. Opt. Soc. Am. A*, **4**, 1543–1553.

SP1695

The Precision of Size Constancy

Suzanne P. McKee*

and

Leslie Welch**

*Smith-Kettlewell Eye Research Institute,

2232 Webster St. San Francisco, CA 94115

and

**University of California at San Diego

La Jolla, CA

Key Words: Objective size, Size constancy, Hyperacuity, Spatial Localization, Stereopsis

ABSTRACT

We compared the precision of objective size judgments, made when target disparity changed at random from trial-to-trial, to angular size judgments made under the same condition. Subjects judged incremental changes in the vertical distance separating a pair of horizontal lines. For the objective judgments (in centimeters), the angle subtended by the target separation decreased with increasing depth consistent with the natural geometry of physical objects. For the angular judgments (in arc minutes), the angular separation did not change with disparity. When the separation subtended an angle less than 10 - 20 arc minutes, objective thresholds were considerably higher than angular thresholds, indicating that size constancy does not function well at small scales. At larger scales ($\geq$20 arc min), angular and objective thresholds were equally precise (~6%) for two of the three subjects. These same two subjects also learned to judge "objective" size when the angular subtense systematically *increased* with increasing depth in an exact inversion of the natural relationship, presumably by changing their response criterion with disparity to minimize error feedback. Although these "anti-constancy" judgments were less precise (~9%) than the constancy judgments, the fact that subjects could learn this task with little practice suggests that constancy itself may be a learned response. Angular thresholds for targets presented only in the fixation plane were significantly lower than the angular thresholds measured with random changes in disparity, indicating the observers do not have direct access to information about angular subtense, independent of target disparity.

INTRODUCTION

Despite the dependence of the size of the retinal image on viewing distance, we can readily judge the physical size of unfamiliar objects over a substantial range of distances -- a phenomenon known as "size constancy". While this ability may seem fairly remarkable, we are, after all, equipped with a number of sensory mechanisms for estimating relative distance. A machine that can measure angular subtense and estimate relative distance can calculate objective size, relative to a reference size viewed at a known distance. In the case of the human observer, this reference size could be a part of the body, like the size of our own hands viewed at arm's length (Morgan, 1989). Since knowledge of objective size is useful for our survival, and our brain has the information needed to calculate objective size, "size constancy" is a fairly predictable aspect of human vision. How well do we make this calculation?

In traditional studies of size constancy, observers were often shown an object at some faraway distance and were asked to adjust the size of an adjacent object until it matched the distant object (Holway and Boring, 1941). Sometimes, the aim of these studies was to determine what was actually seen by the observer -- the objective size or the angular size?. In other cases, the intent was to explore limitations on size constancy, e.g., over what distances observers could match objective size before perhaps regressing to match based on angular size. In one of the most interesting of these studies, Gilinsky (1955) found that observers were able to match either the retinal or the objective size of the test object, depending on the instructions given by the experimenter. Gilinsky's results, subsequently verified in other laboratories (Carlson, 1960; 1977; Leibowitz and Harvey, 1969), indicate that matching is a weak guide to the cognitive (or neural) operations underlying size constancy. One cannot determine whether the observer perceives retinal size, and then corrects this percept by some measure of depth to estimate objective size, or *vice versa.*

There is, however, a psychophysical tool that could reveal the coding sequence. Instead of asking observers what they perceive, we can ask about the *precision* of their judgments of angular and objective size. What is the smallest detectable change in objective size? What is the threshold for discriminating differences in angular size? The precision of psychophysical thresholds is usually limited by noise in the neural pathways coding the stimulus dimensions -- more noise means less precision. If the calculation of objective size involves the simple combination of two independent neural measurements (angular size and depth), objective size judgments should be

consistently less precise than angular size judgments, because the depth measurement should add noise to the calculation; as diagrammed in Figure 1A, there are two sources of noise in the objective size calculation and only one in the angular size calculation[1]. There are, however, coding schemes in which objective size judgments would be *more* precise than angular judgments. For example, the brain may not have direct access to a pure angular or retinal signal (Gogel, 1969), but may instead estimate angular size indirectly by correcting the perceived objective size for an obligatory coupling between size and depth. As shown in Figure 1B, this correction should introduce additional noise, because the noise from the depth estimate enters the calculation twice -- once in the calculation of objective size and again in the indirect calculation of angular size. Either of the models diagrammed in Figure 1 would predict that observers can judge both angular and objective size, because all the information necessary is available for both calculations. The question is which do they do best.

In a recent study, Burbeck (1987a) used a measure of precision, spatial frequency discrimination, to analyze the coding sequence. She compared spatial frequency discrimination for a pair of grating targets presented at a single viewing distance to that for the identical pair of gratings (same object frequencies in cyl/cm) presented at two different distances; she found no significant difference in performance for these two conditions, despite the difference in "retinal spatial frequency" (cyl/deg) necessarily introduced by the second viewing distance. In a second experiment, she observed that observers initially had great difficulty discriminating between targets presented at two different viewing distances on the basis of their "retinal frequencies" (cyl/deg), but were able to learn this "retinal frequency" discrimination with practice. Burbeck concluded that we do *not* have direct access to information about spatial frequency coded in angular units, and must instead estimate angular spatial frequency (and presumably also angular size) indirectly as diagrammed in Figure 1B. Not all angular dimensions are calculated indirectly. McKee and Welch (1989) found that the discrimination of angular velocities was decidedly superior to the discrimination of objective velocities.

Precision is related to measures of variability, so the standard deviations of the matches made in the early constancy studies supply some information about the amount of noise involved in these judgments. Some studies do confirm Burbeck's conclusion. Leibowitz and Harvey (1969) found that the variability of the retinal size matches was greater than the objective size matches: On

[1] We are assuming here that the noise in the neural estimate of depth is roughly comparable, or substantially larger, than the noise in the estimate of angular size. If the noise in the angular size estimate were much larger than in the depth estimate, it would dominate the precision of both types of judgments.

the other hand, Gilinsky's data indicate, that under optimum circumstances (large angular sizes viewed at distances less than 200 feet),the standard deviations of the objective and angular matches were comparable. There are also persistent indications in these studies that size constancy fails at small angular sizes (Boring, 1943; Ross, Jenkins and Johnstone, 1980), although Gibson (1950) believed that size constancy worked at all sizes. Interestingly, in Gibson's study, the mean sizes chosen to match distant objects subtending small angles were close to the true objective size, but the variability of these small-scale matches was much greater than the variability of matches for nearer objects subtending larger angles. This same variability at small angular sizes is apparent in Gilinsky's data, suggesting that the precision of objective size judgments may be significantly degraded at small scales.

These early studies do not provide a clear picture of coding sequence, but that was not their goal. In these studies, the means and standard deviations were based on pooled data, obscuring systematic differences in the data of individual observers about the relative precision of their angular and objective judgments. The measurements were often made at large distances where binocular disparity information was approaching the stereoscopic limit, and the observers were given instructions, but not feedback about the correctness of their matches. With feedback, they might have been to make highly precise judgments that exactly matched either the angular or objective size of the test target.

In the present study, we explored the generality of Burbeck's conclusions for a range of sizes and depths. We gave our subjects the difficult task of abstracting the objective size of a target despite random variations in its apparent depth, but we did provide feedback. Angular size also should be perceived (or perhaps estimated) independent of target depth -- 1 degree equals 1 degree at any depth -- so in a companion study, we measured the ability to judge angular size despite random variations in target depth. We used a very simple, well-studied task to measure the precision of size judgments -- spatial interval discrimination -- in which subjects were required to judge a single dimension, the vertical distance separating a pair of horizontal lines. We manipulated perceived depth by changing only the binocular disparity of the target, but we choose disparities that were easily discriminable and corresponded to physical distances ranging roughly from 1 - 2 meters. Under these severely restricted circumstances, we compared the precision of angular and objective size judgments.

METHODS

We used three different experimental conditions to assess the relative precision of angular and objective size judgments: 1)Incremental judgments of *objective size* with randomly-chosen disparities; 2) Incremental judgments of *angular size* with randomly-chosen disparities; 3)Incremental judgments of size for targets presented only in the *fixation plane* (at one depth)

The target for all conditions is diagrammed in Figure 2; it consisted of two short horizontal lines that defined a vertical separation (or size). Vertical separation was used to minimize the effects of diplopia on the thresholds. The vertical separation was varied parametrically from 0.13 to 2.62 cm for the objective size judgments, and correspondingly, from 3 to 60 arc minutes for the angular size judgments. For a given experimental test run, we chose one particular separation (S), and measured the minimum detectable change in separation (ΔS) using the method of single stimuli. In this method, the subject is shown one of seven possible stimuli chosen from a narrow range, e.g., 2.41, 2.48, 2.55, 2.62, 2.69, 2.76, or 2.83 cm., and is required to judge whether the presented sample is smaller or larger than the mean of the range, equal to 2.62 cm in this example. Thus, the subject judges the test sample against an implicit or remembered standard. While this type of judgment might seem difficult to perform, subjects have no trouble learning the task, and are able to establish an implicit standard with as few as ten practice trials (see Westheimer and McKee, 1977 and McKee, Welch, Taylor and Bowne, 1990 for previous examples of its use). This method had an additional virtue for the present experiments; on our small CRT screens (subtense 3.4 x 4.2 degrees), a visible standard equal to the mean separation (S) and presented adjacent to the test separation ($S \pm \Delta S$), might have introduced confounding cues that would not have had uniform effects in all experimental conditions.

In the objective size condition, the subject's task was to judge small incremental changes in the *objective* size (measured in cm.), despite random variations in disparity. On each trial, the target was presented in one of nine depth planes chosen at random. The nine planes spanned the range from +40 minutes of arc crossed disparity to -40 minutes uncrossed disparity in equally-spaced intervals 10 arc minutes apart; this disparity range corresponded to physical distances ranging from 1.16 to 2.12 meters (see diagram on left side of Fig 3). There is only a small calculated difference between angular and objective size at small disparities ($\pm$ 20 arc min), so, to enhance the difference between these two conditions, we increased the probability that the target

was presented at a large disparity; the target was twice as likely to appear in the four extreme planes ($\pm$ 30 arc min and $\pm$ 40 arc min) as in the fixation plane and the four nearer planes. The targets were actually presented on the screens of an electronic stereoscope at a fixed distance of 1.5 meters, and target disparity was manipulated to produce changes in perceived depth (right side of Fig.3). The mean objective size (in cm) of the vertical separation between the target lines (S) was constant for any experimental test run. When the three-dimensional distance to a real object is increased, the retinal angle subtended by the object necessarily decreases. To simulate the natural relationship between angular subtense and objective size, we varied the angular size presented on the screens of the stereoscope systematically for each depth plane as though the presented objective size ($S \pm \Delta S$) were being viewed at the physical distance specified by the target disparity, i.e. 2.62 cm was set equal to 1.29 degrees at a distance of 1.16 meters (equal to 40 minutes of crossed disparity with respect to the fixation plane), and to 0.71 degrees at a distance of 2.12 meters (40 minutes of uncrossed disparity). Note that the only cue to depth in our experiments was binocular disparity.

The angular size condition was designed to be exactly parallel to the objective size condition. The subject's task was to judge small incremental changes in *angular* size, despite random changes in disparity. On each trial, the target was again presented in one of the nine depth planes chosen at random, but the mean vertical target separation in angular units did *not* vary with target depth. In the third experimental condition, the subject judged small incremental changes in the vertical distance(or size) separating the target lines for targets presented only in the fixation plane.

In a variant of the objective size condition, the *"anti-constancy"* condition, subjects made incremental judgments of size when changes in the angular subtense of the target were completely inverted from the natural arrangement described above for the objective size judgments -- the angular size of the target increased systematically as the apparent physical distance to the target increased. For example, an objective size of 2.62 cm was set equal to 0.71 degrees at a distance of 1.16 meters, and to 1.29 degrees at a distance of 2.12 meters, exactly the opposite of the compensation for apparent physical distance used for objective size condition. Subjects were required to judge incremental changes in "objective" size, while compensating for the inverted changes in angular subtense associated with the depth plane of the target. In short, they were forced to learn a new association between angular subtense and target disparity in order to minimize the error feedback. When the target appeared far away, they judged its size against a large implicit standard, and when it appeared close, they judged its size against a small implicit

standard.

The targets were composed of thin, bright lines, drawn by computer-generated signals on the screens of two Hewlett-Packard 1332A monitors, each equipped with a P-4 phosphor. To insure best performance for larger sizes, e.g. 60 arc min, that necessarily stimulated parafoveal loci, we increased line length with target eccentricity. The length of the target lines was increased parametrically with parametric increases in the vertical distance separating the target lines (S), according to the hyperacuity scaling function originally described by Levi, Klein and Aitsebaomo (1985):

Length =9 (1 + E/.8)

where E is the eccentricity given in degrees, and 9 arc min is the length of the foveal targets. Thus, the horizontal length of the target lines ranged from 9 to 14.6 arc minutes. However, the angular length of the horizontal target lines was not changed with changes in disparity. In a preliminary study, we found that systematic alterations in line length consistent with changes in an objective length did not produce a significant improvement in the objective size increment thresholds possibly because target length was obscured by diplopia at the extreme disparities.

The images on the two monitors were superimposed by a beam-splitting pellicle. Orthogonally-oriented polarizers placed in front of the monitors and the subject's eyes insured that only one screen was visible to each eye. Before each test stimulus, a fixation pattern was presented for 800 msec, and then turned off just before the presentation of the test target. The fixation pattern, presented binocularly, consisted of four corner brackets forming an implicit square 60 arc minutes on a side with a bright point in the center of the square. The lines forming the test configuration were presented symmetrically arrayed around the center of the fixation pattern. As the fixation pattern might have provided an additional alignment cue for the largest angular test size (60 arc min), it was increased to 90 arc min for that size alone. The target duration for most of these experiments was 150 msec, too brief to permit a voluntary shift in convergence. Since the targets were presented at disparities symmetrically arrayed around the fixation plane, e.g. $\pm$ 40 arc minutes, we may assume that the subjects kept their eyes close to the fixation plane for most of the experimental session, and that additional proprioceptive cues from convergence and accommodation to depth planes other than the fixation plane were not available at these brief durations. For some of the experiments, duration was increased to either 1000 or 1500 msec, a time sufficient to allow convergence to the depth plane occupied by the target. If convergence were accompanied by accommodation, as is probable, the targets off the fixation plane would have been slightly out of focus in this long duration condition (maximally ~1/3 diopter).

Line luminance was measured with a Pritchard photometer; the test pattern for these measurements was a long vertical line created with the same timing and intensity characteristics as the experimental target lines. The measurements for the line on each screen were made in the dark, through the pellicle, at two different distances (147 cm and 258 cm) and with two different apertures (1 deg and 20 arc min) for accuracy. The photometer measurements (which are measured in cd/m^2) were converted to *line* luminances using the formula:

$$L = (\text{meter reading})\ \pi\ d/4$$

where d is the diameter of the measuring aperture. The line on each screen measured approximately 0.022 cd/m. It should be noted that these small *line* intensities produce bright, easily visible lines; if we filled the screen with such lines, one line every 1mm, the mean luminance of the screen would equal 22 cd/m^2. The background luminance measured 0.52 cd/m^2. Overhead fluorescent lighting located about 2 meters from the CRT monitors supplied indirect illumination of the experimental setting at a moderate photopic level. Room furniture and experimental equipment were clearly visible. The only immediate reference frame for the targets was the 14 x 11.5 cm opening in the cube that contained the beamsplitting pellicle; the opening was 45 cm in front of the fixation plane. The edges of the CRT screens and the pellicle were very dim and partially obscured by the cube housing the pellicle. Moreover, the two CRT screens were not superimposed optically, so they appeared to float at an indistinct distance. The only target providing good information for convergence in the fixation plane was the bright square fixation pattern presented before each test target. In a control experiment for our earlier study of constancy (McKee and Welch, 1989), we found that making these measurements in total darkness did not change the pattern of results from those obtained when room furniture was visible

Each of the increment thresholds presented in this paper is based on at least 300 trials, usually from two separate test runs of 150 trials each. Additional experimental sets were taken for some sizes when there was a substantial difference between the thresholds for the first and second sets; all sets were summed to estimate the threshold. The thresholds were estimated from the psychometric functions generated by plotting the percentage of trials on which the observer responded that the presented separation ($S \pm \Delta S$) was larger than the mean separation (S) as a function of the distance separating the target lines. A cumulative normal function was fitted to the resulting function by probit analysis. Threshold was defined as that incremental change in size that produced a change in response rate from the 50% to the 75% level, equal to a d' of 0.675. Each experimental test run began with 10-20 practice trials. All measurements were made using audible

error feedback.

For those cases where we wished to measure the perceived size as a function of target disparity, data from 900 - 1000 trials were accumulated at a single criterion size, e.g 1.2 cm $\pm \Delta$ for the objective size judgments or 30 arc min $\pm \Delta$ for the angular size judgments, and sorted by disparity into separate bins. A Point of Subjective Equality (stimulus corresponding to the 50% response value on the psychometric function) were estimated from the psychometric function generated by the data in each of the separate bins. The standard errors of the PSE's hovered between 1 and 2 %, except for one case, the "anti-constancy" PSE corresponding to a disparity of 10 arc min, where erratic performance and a small number of trials combined to make the PSE indeterminate. An increment threshold was also estimated from the pooled data for all disparities.

The two authors and a third experienced observer served as subjects for these experiments; all three had 20/20 visual acuity for the viewing distance of 1.5 meters and good stereoacuity, and all had had much practice on size or separation judgments in previous studies. Our original intent was to assess the natural capabilities of well-trained adult observers to judge angular and objective sizes. These experiments were *not* designed to measure the ability to learn novel stimulus criteria. Subjects LW and SM spent a brief period (roughly 300 trials) practicing with both types of judgments before the data presented here were collected; neither author noted great difficulty in performing either task, but both had participated in an earlier study in which judgments of angular and objective size for a single size had been a control condition for velocity constancy (McKee and Welch, 1989). The main function of practice was to reduce confusion between the "objective size" conditions, and the "angular size" conditions. Subjects tended to take "objective" thresholds, and "angular" thresholds in blocks, although sometimes they interspersed conditions as needed to complete at least 300 trials for all tested sizes in all experimental conditions. When switching from one condition to another, a subject would often take some practice (40 - 150 trials) at one size to be reminded of the appropriate criterion for that block of thresholds; data from these designated practice sets were discarded. Questions arising from our initial results lead us to design the anti-constancy experiment. Clearly, "anti-constancy" is not a natural part of human experience, so for these measurements, subjects LW and SM practiced for 600 trials before collecting the data shown in Figures 10 and 11. The average Weber fractions from the practice sets were 0.125 for LW and 0.10 for SM -- slightly higher than the Weber fractions for "anti-constancy" shown in Figure 11 which were based on accumulated data from several runs (900 trials for LW and 1050 trials for SM), taken after completing the 600 practice trials. These thresholds do not necessarily represent

asymptotic performance; with continued practice, performance may have improved.

The third subject (WAM) did not seem to have any natural ability to estimate objective size from disparity at a short duration (150 msec) despite abundant practice. He felt that he might be able to perform this task better at a longer duration, so we repeated these size measurements at a duration of 1000 msec. The data presented in Figure 5 were taken following several thousand practice trials on objective and angular size judgments at the shorter duration. The results for the 1500 msec duration shown in Figure 8 were taken after additional specific practice on both objective and angular judgments at a size of 1.3 cm and 30 arc min. For those interested in the genetic contribution to these perceptual abilities, subject WAM shares half his genetic endowment with subject SM; his father declined to participate in these experiments.

RESULTS

Our first experiment compared objective size judgments, made when target disparity changed randomly from trial-to-trial, to angular size judgments made under the same condition. The resulting increment thresholds, each based on the pooled data from all nine depth planes (300 trials), are shown in Figure 4 for a range of sizes. Thresholds for targets presented only in the fixation plane are also plotted in the same figure. For ease of comparison, all thresholds are presented in a common framework based on angular units. The separation plotted on the horizontal axis refers to the mean angular size. By design, the angle subtended by the mean objective size at 1.5 meters (the fixation plane) is equal to this mean angular size. The minimum detectable increment in objective units (cm) was first translated into the dimensionless Weber fraction, and then plotted as an incremental change in the mean angular size corresponding to the mean objective size.

Two things are apparent from Figure 4. First, while the objective size thresholds are decidedly less precise than the angular size thresholds at small separations ($\leq$ 10 arc minutes), they are nearly equal to the angular thresholds at separations greater than 10 arc minutes. Second, the random disparity angular thresholds are much less precise than thresholds measured in a single depth plane, indicating that, when we view a target binocularly, we do not have access to a pure angular or retinal estimate of size uncontaminated by depth signals

Size Constancy at Small Angular Subtense

Size constancy does not operate efficiently for separations subtending small angles. In

figure 5, we have plotted the ratio of objective to angular thresholds as a function of mean separation. The ratio declines steadily, reaching a value close to 1 between 10 and 20 arc minutes. A third subject, WAM, had great difficulty judging objective size at the short duration used for the other two subjects, so he repeated the measurements for a longer duration of 1 second. The ratio of his thresholds show the same pattern as that of the other two subjects.

Increment judgments at very small angular sizes are known to be relatively less precise than judgments at larger angular sizes. The Weber fraction for judging changes in a 1 min spatial interval is 0.2 - 0.3, while the Weber fraction for a 10 min spatial interval is about 0.03 (Westheimer and McKee, 1977; Klein and Levi, 1987). Our measurements of objective size thresholds are necessarily based on data from a range of angular sizes, consistent with changes in the depth plane, e.g.,the angle subtended by 0.2 cm (5 arc minutes at a viewing distance of 1.5 meters) ranges from 3.5 arc min to 6.5 arc min for the disparities used here. The objective size judgments for small scales might be elevated because they are based on a mixture of precise and imprecise angular signals. However, as long as we operate within the range of angular sizes where Weber's Law holds, our objective thresholds should not be affected by this pooling of large and small angular signals. As an example, consider the angular sizes associated with the mean objective size of 1.2 cm; the angular sizes range roughly from 20 to 40 arc minutes, so if the Weber fraction were 0.03, the increment thresholds should range from 0.6 to 1.2 arc minutes producing an average increment threshold of 0.9 arc minutes. At a distance of 1.5 meters, 1.2 cm subtends 30 arc minutes so the average increment threshold of 0.9 minutes is equal to a Weber fraction of 0.03, just what we would predict for the threshold corresponding to the mean of this range. For these two subjects, Weber's Law holds for separations as small as 3 arc minutes (see the fixation plane data in Figure 4). Therefore, the differences between the angular and objective thresholds at 5, 8 or 10 minutes cannot easily be explained by this mixture argument. To check this conclusion, we increased the number of trials for two separations, 5 arc min (or 0.2 cm) and 30 arc min (or 1.2 cm), thereby obtaining adequate estimates of the perceived mean sizes (PSE) and increment thresholds associated with each of the nine depth planes. At the small objective size (0.2 cm), the Weber fractions ($\Delta s/s$) for the separate depth planes showed no significant trend with increasing depth (decreasing angular size) -- further evidence against the mixture argument.

In Figure 6, we have plotted the perceived mean sizes (PSE's), once again in a common framework based on angular units. If the subjects could scale their judgments appropriately for depth, the objective means in *angular* units should increase systematically with decreasing depth.

falling on the diagonal line drawn in each graph. The angular means should, of course, remain constant and fall along the horizontal line also drawn in each graph. Indeed, all of the angular means follow the predicted horizontal line, although there is a hint of interference from "size constancy" in the 30 min data of subject SM; the mean angular size looks larger to her at the farthest depth plane than at the nearest plane. Both subjects do a fair job of responding to objective size at the larger scale; the means fall close to the oblique line in agreement with prediction. Neither subject could judge objective size appropriately for the target subtending the small angle. The means for the objective and angular size judgments overlap at all except the most extreme disparities, where the objective means separate from the angular means, perhaps reflecting a half-hearted attempt by the subjects to respond in a manner consistent with the feedback reinforcing objective size judgments.[2] We conclude that objective size thresholds are elevated at separations below about 20 arc minutes because the subjects were unable to take the depth fully into account in responding to these small sizes.

Random Changes in Disparity Degrade Angular Judgments

At larger sizes, angular and objective thresholds are equally precise, but both are significantly higher than the thresholds for targets presented only in the fixation plane. As Figure 7 shows, the ratio of the random disparity *angular* thresholds to the fixation plane thresholds is nearly a factor of 2 at all separations, indicating that the noise elevating the random disparity thresholds is multiplicative; it doubles the threshold over the whole range independent of the absolute magnitude of the increment threshold. An observer can judge angular size better if he or she keeps one eye closed, obscuring the random shifts in disparity. For example, we measured a "random disparity" angular threshold with one screen covered so that the target (one stereo half-image) shifted left or right from trial-to-trial over a 40 minute range; the threshold measured with these random lateral displacements was identical to the threshold for the fixation plane condition where the target was presented in one position. Thus, changes in depth, and not simply changes in retinal position, are responsible for the elevation in the thresholds.

Disparity could, in itself, increase the noise, because the neural units that code large

[2] An ANOVA showed that the PSE's for the small objective size were significantly different ($p=0.02$) from the PSE's for the small angular size provided that the data from the extreme disparities were included in the analysis. When the data from $\pm$ 40 min disparities were removed from the analysis, the difference between the angular and objective PSE's for the 5 min size was no longer significant ($p = 0.3$). At the larger 30 min size, the PSE's for objective size were significantly different from the PSE's for angular size with or without the extreme disparities ($p= 0.0001$).

disparities may be much coarser (larger receptive fields) than the units that code small disparities (Marr and Poggio, 1979). Large receptive fields are spatially extended and have shallow weighting functions, so their signals provide less precise information about location than smaller receptive fields. If spatial localization of the horizontal target lines is mediated by these coarser disparity units, the random disparity thresholds would necessarily be less precise than the fixation plane thresholds. Why would the brain use the signals from these coarse disparity units to encode position when more precise smaller units are available? Here we must assume that position and disparity are jointly encoded, that there is an obligatory coupling between depth and location. To register the disparity of the target, an individual with normal stereopsis must use the coarser disparity units at a cost in the precision of spatial location. This coupling between disparity and location has been observed in another context; McKee, Levi and Bowne (1990) found that the precise signal associated with a monocular vernier target was obscured when the vernier target was paired stereoscopically with a disparate target in the other eye. Nevertheless, this disparity explanation is not completely satisfactory because the effect of these coarser disparity units should be similar to the effects of eccentricity on spatial localization -- there should be a significant elevation in the thresholds for small separations, but no effect on the thresholds for larger separations (Yap, Levi and Klein, 1987). As Figure 7 shows, randomly varying the disparity has a fairly uniform effect on the angular thresholds for all separations.

An alternative explanation for the noise in the random disparity thresholds is reference uncertainty. White, Levi and Aitsebaomo (1991) have argued that reference uncertainty has a multiplicative effect on the increment thresholds for separation, for targets presented in the fixation plane (see also Klein and Levi, 1987, Morgan, 1991). It is easy to see why this source of noise is multiplicative. The subjects make these judgments using an implicit reference that they probably estimate by taking a running average of the preceding three or four trials. Even if they were perfect at this averaging process, the average would often be either too small or too large with the result that the stimulus on the current trial could be "perfectly" encoded, but still judged incorrectly. Usually, in the Method of Constant Stimuli, the size of the incremental steps are increased proportionally as the size of the mean separation increases since the threshold also grows proportionally, i.e., Weber's law holds over a large range of separations. Thus, the errors produced by reference uncertainty are necessarily magnified by the step size, producing a multiplicative effect on the thresholds. Why do random variations in disparity increase this reference uncertainty? If the subjects were attempting to apply a different angular reference to

every depth plane --use multiple standards -- as scaling models of size constancy imply (Andrews, 1964; Morgan, 1991), they might misjudge the depth plane of the target, or estimate the implicit reference by averaging trials from adjacent depth planes, thereby increasing the reference error. The problem with this explanation is that there is no reason for the subjects to use multiple standards in judging *angular* size. The angular separation between the target lines is not changing with changes in disparity, so the subjects can employ a single implicit reference in judging the angular distance separating the lines.

To distinguish between noise attributable to disparity *per se* and noise introduced by reference uncertainty, we increased target duration to 1500 msec and asked subjects to converge to the plane of the target on every trial. Under these conditions, we repeated the PSE and threshold measurements for one separation (1.3 cm and 30 arc minutes). This task is very demanding -- a bit like doing oculomotor push-ups -- and subject LW was unable to perform the task because it induced severe headaches. The data from the other two subjects are presented in Figure 8. Increasing duration has improved performance for both subjects. For subject WAM, the PSE's for angular and objective size are completely intermixed at the short duration, but at the longer duration, his PSE's are closer to the predicted functions resulting in an improvement in both thresholds, as shown by his Weber fractions (see boxes at the bottom of each graph). Nevertheless, his objective size Weber fraction is higher than his angular size Weber fraction[3], and his angular Weber fraction is still significantly higher than his fixation plane Weber fraction, even at the 1500 msec duration. For subject SM, the objective and angular Weber fractions are exactly equal at the longer duration, but both are stubbornly higher than the fixation plane Weber fraction despite the reduction in target disparity produced by converging to the plane of the target.

What accounts for this persistent decrement in the precision of the random disparity angular judgments? Convergence may not have been perfect, leaving the target with a small uncorrected disparity, and it may also have induced changes in accommodation, thereby introducing a slight amount of blur in the target (maximally 1/3 of a diopter). However, thresholds for a 30 min separation are not likely to be degraded by a minor amount of disparity or defocus (Klein and Levi, 1990b). We suspect that reference uncertainty is a source of noise for the angular thresholds as well as for the objective thresholds. Since reference uncertainty presumably depends on the number of simultaneously-held references, our thresholds should improve if we reduced the

[3] The sharp-eyed reader may discover that the ratio of angular to objective thresholds for subject WAM is higher in this figure than in Figure 5. The data in figure 5 were taken four months earlier. Upon return to work, subject WAM was given several additional days of practice before we made the measurements shown in Figure 8. Nevertheless, some "backsliding" is evident.

number of tested depth planes from nine to two. In a final appraisal of the contribution of reference uncertainty to angular judgments, we compared the effects of reducing the number of depth planes to the effects of reducing target disparity for one subject (SM). The thresholds, based on the averaged data from many interspersed measurements, are presented in Figure 9 for each of twelve conditions. In the "Two Planes" condition, the target was randomly presented either in the fixation plane or at a disparity of 30 arc minutes, a technique that forced the subject to shift convergence repeatedly as in the "Nine Planes" condition. The target was always presented in the fixation plane for the "One Plane" condition, and the "Nine Planes" condition was identical to the random disparity condition described above.

The data in Figure 9 show that the effects of disparity and reference uncertainty depend on the size of the mean separation. For the small 5 min separation, disparity is solely responsible for degrading the precision of the angular threshold. At the short duration (150 msec), the 5 min threshold is equally elevated by disparity whether it is associated with two planes or nine. If the subject is given time to converge (1500 msec), there is no significant difference between the threshold for the fixation plane ("One Plane") and the threshold for the random disparity condition ("Nine Planes"). Compare these results to those for the larger 30 min separation where increasing the duration had almost no effect, but increasing the number of depth planes -- increasing the reference uncertainty -- had a much greater effect. In the "Two Planes" condition, the target was presented with a large disparity on half the trials, but the threshold for this condition is the same as the "One Plane" condition for both durations, a result that shows that disparity adds little noise to these large-scale thresholds. Disparity and reference uncertainty both degrade angular judgments, but the two sources of noise operate at different spatial scales, conspiring to produce a nearly-uniform multiplicative effect on the angular thresholds. While it is reasonable that the "neural" blur produced by large disparities would have its greatest effect on the tiny incremental thresholds associated with small separations (thresholds in the hyperacuity range), the detrimental effect of reference uncertainty at large angular sizes is less easily explained, and will be discussed later.

"Anti-Constancy" vs. Constancy

We were surprised to discover that subject WAM had difficulty judging objective size, so, for demonstration purposes, we constructed a target consisting of two identical pairs of lines, each pair separated by the same angular separation. Both pairs were presented simultaneously, but one pair was shown with an uncrossed disparity of 10 arc minutes. For subject SM, the separation

between the distant pair looked larger than the separation of the nearer pair, but to subject WAM, the separation between both pairs looked identical, even given time for scrutiny. Apparently, he has no perceptual size constancy based on disparity alone. How then was he able to make the objective size judgments? He said that he had learned to use different references for different distances to minimize error feedback. For this subject, size constancy was not a natural response induced by disparity, but was instead an arbitrary recalibration.

Could the other two subjects learn to respond on some completely arbitrary basis? To answer this question, we created an experimental condition in which the angular size of the target *increased* with increasing distance, exactly inverting the natural relationship, a condition we labeled "anti-constancy". As in the previous experiment on size constancy, subjects were required to judge target size while compensating for an orderly change in angular subtense with disparity, but in this case, their responses were judged according to criteria that had no natural counterparts. Error feedback was given throughout the experiment to assist the subjects in establishing an appropriate set of references for all depth planes. Following a small amount of practice, subjects LW and SM collected enough data (>900 trials) to estimate "anti-constancy" PSE's for single mean size (1.3 cm equal to 30 arc min at 1.5 meters) at each of the nine depth planes,. These PSE's are plotted in Figure 10, along with the PSE's for normal constancy, redrawn from Figure 6. Although both authors were fairly good at making these "anti-constancy" judgments, neither perceived the target as being of a constant size. On the contrary, the most distant target appeared enormous, jointly magnified by the increased angular size programmed into our display, and by size constancy -- our well-learned tendency to perceive the size of more distant objects as larger than nearer objects. We both adopted a similar strategy for making these judgments. We imagined a three-dimensional wedge with the small end pointed towards us, and, as each target was briefly displayed, we judged the target against this imagined reference frame -- a visualization of the re-scaling needed to reduce the number of errors signaled by the computer.

The judgments for "anti-constancy" are somewhat less precise (roughly a factor of 1.5) than for constancy. As Figure 11 shows, subjects LW and SM could respond to the inverted "anti-constancy" relationship with the about same precision that subject WAM could respond to the normal "constancy" relationship (lower right side of Figure 8). This result is especially curious considering that the duration used for the anti-constancy thresholds shown in Figure 11 was 150 msec, one tenth of duration used for the constancy judgments of subject WAM shown in Figure 8. The ease with which these two subjects learned this unnatural trick suggests that natural size constancy may also be learned response (Helmholtz, 1868; Morgan, 1991).

DISCUSSION

In our artificial viewing conditions, the best subjects could judge objective size with a precision of 5 - 6 %. These low thresholds, based on a liberal psychophysical criterion (d'=.675), may nonetheless underestimate the precision of size constancy in natural surroundings, where the rich array of visual information could substantially reduce uncertainty about depth and size. Thus, we were pleased to discover that the most precise Weber fractions (standard deviation/mean size -- d' = 1.0) derived from Gilinsky's data were between 6 - 7%, since Gilinsky performed her experiments in an open field and allowed her subjects unlimited viewing time. Taken as a whole, our results lead to two rather puzzling conclusions. First, observers cannot judge objective size (cm) with the same precision as they judge angular size (deg) for targets subtending less than 10 arc min, but, beyond this range, some observers can judge objective and angular size with equal proficiency. Second, observers do not have direct access to information about angular subtense, at least not for targets viewed binocularly. These conclusions are supported by earlier work on object perception and size constancy (Boring, 1943; Gilinsky, 1955; Gogel, 1969; Ross, Jenkins and Johnstone, 1980; Burbeck, 1987a), so it is worth speculating about the reasons for these limitations on the processing of visual size.

Why should there be a lower limit on size constancy? There is increasing evidence that different cues are used for judging small distances than for judging large ones (Burbeck, 1987b; Klein and Levi, 1985; 1987; Levi, Klein and Yap, 1988; Levi and Klein, 1990a; Wilson, 1991). The cues at small distances are more akin to contrast judgments than to distance judgments, in that observers use subtle changes in the light distribution to detect changes in position. To take an extreme example, consider a pair of bright lines separated by a barely discernible gap of 2 arc minutes. If the separation is decreased, the lines fuse into a single bright bar; if the separation is increased, the gap becomes darker and more distinct. At slightly larger distances, e.g., 5 arc minutes, the apparent width of the central dark gap can be compared to the apparent width of either of the bright target lines. While these particular cues are specific to our task, thresholds for judging the dimensions of features that subtend small angles (< than 10 arc min) are probably limited by noise in mechanisms that code local changes in contrast[4].

[4] Note that observers can easily discriminate between changes in the overall luminance or contrast of the target, and changes in its size or the distance separating the lines. Observers are responding to delicate changes in the *shape* of the light distribution to judge these small changes in position, but their ability to detect these changes is thought to be limited by contrast sensitivity.

In the fovea, increment thresholds for small distances are typically in the hyperacuity range, i.e., smaller than the size of a single cone. Hyperacuity models have been moderately successful at predicting these small spatial thresholds from the contrast-driven responses of different-sized spatial filters -- the mathematical representations of simple cortical units (Klein and Levi, 1985; Wilson, 1986). Disparity degrades these hyperacuity thresholds (see Figure 9), presumably because the coarse spatial filters that respond to large disparities(Tyler, 1975; Marr and Poggio, 1979) effectively blur subtle changes in the light distribution. However, target disparity was varied randomly for both the angular and objective judgments, yet, at small separations, the objective judgments were significantly less precise that the angular ones. The reason is fairly obvious. In the objective condition, the angular separation between the targets changes with disparity completely obscuring the subtle changes the observer uses to make these fine spatial judgments. In order to use these contrast-dependent cues to judge objective size, the observer would need a separate template of the light distribution of the reference size for every tested disparity. An ideal observer could readily store these multiple templates, but the real observer has many opportunities to make mistakes -- to misjudge the disparity or to use the wrong template -- with a resulting decrement in performance.

Doesn't the real observer have this problem with objective judgments at all sizes? Why should this problem disappear at larger scales? For one thing, the observer uses a different strategy to judge the distance between widely-separated features. When there is no overlap between the retinal images of the target lines, the light distribution provides little information about separation, so instead each target line is separately localized according to its spatial co-ordinates ("its local sign"). The evidence for this premise is that, once target separation exceeds 5 - 10 arc minutes, thresholds are not affected by target contrast, by target spatial frequency content, or even by whether the targets are of opposite contrast sign (Morgan and Regan, 1987; Burbeck, 1987b, Levi and Westheimer, 1987). At small scales, the contrast-dependent changes in the light distribution provide more precise information about separation than the spatial co-ordinates of the target lines. At large scales, this contrast-dependent information is no longer available so separation is economically coded as the distance between the spatial co-ordinates corresponding to the "local signs". Several studies have concluded that the "local sign" of a single feature in the fovea is known to the precision specified by a single cone (Zeevi and Mangoubi, 1984; McKee, Welch, Taylor and Bowne, 1990; White *et al*,1991) Interestingly, in our measurements, size constancy fails as the increment threshold approaches the size of a foveal cone. We speculate that size constancy operates only on the information supplied by the "local signs", perhaps by re-

scaling the distance between the co-ordinates of each individual contour with changes in depth (Andrews, 1964), or perhaps by maintaining a separate reference size for every disparity (Morgan, 1991). Certainly it would require less information (fewer "bits") to store a neural representation of the distance between the spatial co-ordinates than to store a complete template of the luminance distribution of the reference size for every tested disparity.

On which spatial map do these co-ordinates lie? Traditionally, "local signs" are thought to refer to retinal co-ordinates, but our data indicate that an observer with normal stereopsis viewing a target binocularly does not have access to retinal or monocular information (see also McKee, Levi and Bowne, 1990). Else there would be no discrepancy between the fixation plane thresholds and the angular thresholds measured with random changes in disparity, a discrepancy that persists at large scales, even when the observer is given time to converge on the target plane. The physical information available at the eye is angular subtense, but it is unlikely that information about angular subtense or retinal location is preserved beyond the binocular confluence occurring at striate cortex. The first primitive map of location may be generated at the striate level, but it already represents a transformation of the retinal co-ordinates (Levi *et al*, 1985). For off-horopter targets, perceived location necessarily reflects signals from both retinae, and is not congruent with either (Nelson, 1977; Sheedy and Fry, 1979; Rose and Blake, 1988). Off-horopter neural units are undoubtedly noisier than fixation plane units (McKee, Levi and Bowne, 1990), but if this were the only additional source of noise limiting the precision of angular size judgments, then allowing subjects sufficient time to change convergence should make the angular thresholds for multiple planes equal to the threshold for the fixation plane. The persisting decrement in the large scale thresholds suggests that there is an obligatory coupling between depth and size, so that spatial co-ordinates are specified in three dimensions, not two.

Does this result imply that the natural metric of our nervous system is based on object space? While our data do not support a metric based on angular subtense, they also do not support a metric based on objective size. Two of the subjects -- the authors --had a strong illusion of size constancy induced by target disparity, but their objective thresholds were not better than their angular thresholds, and the objective thresholds of the third subject were decisively worse than his angular thresholds. Unlike the results of Burbeck 's study (1987a), our object-based thresholds were never equal to thresholds for targets presented in a single plane. In Burbeck's study, subjects discriminated between the reference spatial frequency presented at one physical distance and the test frequency presented at another, with the test and reference exchanging positions at random

from trial-to-trial. Perhaps the presence of a simultaneously-visible reference or the additional cues to depth facilitated the recalibration for objective size. However, we think the most important difference between her study and our own is that her observers only had to cope with two planes, whereas our observers were dealing with nine.

Morgan (1991) has argued persuasively that size constancy reflects the human ability to use multiple references in judging size. In the case of size constancy, target disparity acts as the cue to the appropriate reference. For example, our observers may have used a larger angular reference when the target was in the forward planes than when the target was in the rear planes. Our "anti-constancy" results show that a subject can quickly learn an orderly, but arbitrary, reference system in judging size. Morgan (1991) asked subjects to judge the width separating a pair of lines presented at one of four different orientations on interspersed trials; there was a different implicit reference width for each orientation, so subjects were forced to switch their size criterion with changes in target orientation. He found that the precision of size judgments was unaffected when the size criterion changed systematically with changes in orientation or position, for four different reference sizes. A slight, but perceptible, loss in precision is apparent in his data as the number of references increased from four to eight; Weber fractions for his best subject increased from about 3% to about 5%. This small increase in thresholds is comparable to the difference between our single plane, single criterion thresholds, and our multi-plane, multi-criteria thresholds (1500 msec) for both objective and angular size judgments.

Angular size? Why would the observer have needed a reference to judge angular size, since angular size did not change with disparity? Strange as it seems, the angular size judgments seem to suffer from the same uncertainty as the objective size judgments. We can only speculate that this uncertainty is due to an interaction between two sources of depth information. Far from being only the "raw data" of experience, angular size is the basis of a powerful cue to depth -- size or texture gradients (Gibson, 1950). The depth signalled by angular size is necessarily ambiguous, because the features subtending a smaller angle are either farther away, or physically smaller. Normally, we view the world with both eyes open, so we can use relative disparity information to check whether the decrease in angular subtense is consistent with the apparent depth. What happens if there is an inconsistency? If the texture were continuous, or if the features were familiar, angular size information might carry more weight than disparity (Maloney and Landy, 1989), but in our spartan displays, relative disparity is a better indicator of depth. Apparently the coupling between depth and angular subtense creates a slight uncertainty about whether the presented stimulus is

smaller or larger than the reference stimulus. In short, the diagrams in Figure 1 are too simplistic. Constancy may be well-learned reference system, in which the co-ordinates are re-scaled with depth, but this rescaling process can interfere with direct estimates of angular size.

In Figure 12, we have summarized our conclusions in two flow charts. At small scales, the observer can use one of two cues to detect changes in angular size or distance -- the shape of the light distribution, or the distance between the spatial co-ordinates corresponding to the dimensions of the target. Subtle shifts in the light distribution provide the most precise signals for making angular judgments, although these signals can be degraded by disparity (the large noise in the off-horopter units). For judgments of objective size, this entire light distribution cannot be rescaled for different depth planes without significant loss of information, so, instead, the scaled distance cue becomes the better source of information about objective size. At large scales, only the distance between spatial co-ordinates is used for either angular or objective size judgments. This information is nearly immune to noise in the off-horopter binocular units, but both types of judgment suffer some sligh: loss of precision because of the well-learned relationship between depth and size.

ACKNOWLEDGMENTS

This research was supported by the Air Force Office of Scientific Research Grant #89-0035 and NIH Core Grant 5 P-30-EY-01186. We thank Professor Michael J. Morgan for his thoughtful critique of the manuscript, Dr. Samuel Bowne for helpful discussions and Doug Taylor for his carefully conceived programs and technical assistance.

REFERENCES

Andrews D.P. (1964) Error-correcting perceptual mechanisms. Quarterly Journal of Experimental Psychology, 16, 105-115.

Boring, E.G. (1943) The moon illusion. American Journal of Physics, 11, 55-60.

Burbeck, C.A. (1987a) On the locus of spatial frequency discrimination. Journal of the Optical Society of America..A. 4, 1807-1813.

Burbeck, C.A. (1987b) Position and spatial frequency in large-scale localization judgments. Vision Research , 27, 417-427.

Carlson, V.R. (1960) Overestimation in size constancy American. Journal of. Psychology, 73, 199-213.

Carlson, V.R. (1977) Instructions and perceptual constancy judgments. In, Stability and Constancy in Visual Perception: Mechanisms and Processes, edited by W. Epstein, New York: John Wiley and Sons. 217-254.

Gibson, J.J. (1950) The Perception of the Visual World. Boston: Houghton-Mifflin.

Gilinsky, A.S. (1955) The effect of attitude upon the perception of size. American. Journal of. Psychology, 68, 173-192.

Gogel, W.C. (1969) The sensing of retinal size. Vision Research., 9, 1079-1094

Helmholtz, H.V. (1868) The recent progress of the theory of vision. Trans. by P.H. Pye-Smith. In, Popular Scientific Lectures, edited by M. Kline, New York: Dover Publications, 1982

Holway, A.H. and Boring, E.G. (1941) Determinants of apparent visual size with distance variant. American. Journal of. Psychology, 54, 21-37.

Klein, S.A. & Levi, D.M (1985) Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. Journal of the Optical Society of America..A.2, 1170-1190.

Klein, S.A. & Levi, D.M.(1987) Position sense of the peripheral retina. Journal of the Optical Society of America..A. 4.1543-1553.

Leibowitz, H.W. and Harvey, L.W. (1969) Effect of instructions, environment, and type of test object on matched size. Journal. of Experimental Psycholgy., 81, 36-43.

Levi, D. M., Klein, S.A. & Aitsebaomo, P. (1985) Vernier acuity, crowding and cortical magnification. Vision Research, 25, 963-977.

Levi, D.M., Klein, S.A. & Yap, Y.L. (1988) "Weber's Law for position: unconfounding the role of separation and eccentricity.Vision Research,28 597-603

Levi, D.M. and Klein, S.A. (1990a) The role of separation and eccentricity in encoding position. Vision Research, 30 557-586.

Levi, D.M. and Klein, S.A. (1990b) Equivalent intrinsic blur in spatial vision. Vision Research, 30. 1971-1993.

Levi, D.M. and Westheimer, G. (1987) Spatial-interval discrimination in the human fovea: what delimits the interval? Journal of the Optical Society of America, A, 4, 1304-1313.

Maloney, L.T. and Landy, M.S. (1989) A statistical framework for robust fusion of depth information. Proceedings of SPIE: Visual Communications and Image Processing, Part 2, 11543-1163.

Marr, D. & Poggio, T. (1979) A computational theory of human stereoscopic vision. Proceedings of the Royal Society, B204, 301-328.

McKee, S.P., Levi, D.M. & Bowne, S.F. (1990)., The imprecision of stereopsis. Vision Research, 30, 1763-1779

McKee, S P. & Welch, L. (1989) Is there a constancy for velocity? Vision Research, 29, 553-561.

McKee, S.P., Welch, L., Taylor, D.G. & Bowne, S.F. (1990) Finding the common bond: Stereoacuity and the other hyperacuities. Vision Research, 30, 879-891.

Morgan, M.J.(1989) Vision of solid objects. Nature, 339, 101-103.

Morgan, M.J. (1991) Hyperacuity. In Regan, D. (Ed.) Spatial Vision . London: Macmillan, in press.

Morgan, M.J. (1991) On the scaling of size judgments by angular cues. Submitted to Vision Research.

Morgan, M.J. & Regan, D. (1987) Opponent model for line interval discrimination: interval and vernier performance compared. Vision Research 27 107-118.

Nelson, J.I. (1975) Globality and stereoscopic fusion in binocular vision. Journal of Theoretical Biology, 49, 1-88.

Rose, D. & Blake, R. (1988) Mislocalization of diplopic images. Journal of the Optical Socity of America A, 5 1512-1521

Ross, J., Jenkins, B. & Johnstone, J.R. (1980) Size constancy fails below half a degree. Nature, 283, 473-474.

Sheedy, J.E. & Fry, G. (1979) The perceived direction of the binocular image. Vision Research, 19, 201-211.

Tyler, C.W. (1975) Spatial organization of binocular disparity sensitivity. Vision Research,15, 583-590.

Westheimer, G. & McKee, S.P. (1977) Spatial configurations for visual hyperacuity. Vision Research, 17, 941-947.

White, J.M., Levi, D.M. & Aitsebaomo, A.P. (1991) Spatial localization without visual references. Submitted to Vision Research.

Wilson, H. R. (1986) Responses of spatial mechanisms can explain hyperacuity. Vision Research, 26, 453-469.

Wilson, H. R. (1991) Psychophysical models of spatial vision and hyperacuity. In Regan, D. (Ed.) Spatial Vision . London: Macmillan, in press.

Yap, Y.L., Levi, D.M. & Klein, S.A (1987) Peripheral hyperacuity: 3-Dot bisection scales to a single factor from 0 to 10 deg. Journal of the Optical Socity of America, A 4, 1554-1561.

Zeevi, Y.Y. & Mangoubi, S.S. (1984) Vernier acuity with noisy lines: estimation of relative position uncertainty. Biological Cybernetics, 50, 371-376

FIGURE LEGENDS

Figure 1

Flow chart diagramming sources in noise that limit the precision of objective and angular judgments. In Figure 1A, observer has direct access to information about angular subtense. Objective size is calculated from depth and angular size, and the additional noise from the depth estimate makes objective judgments less precise than angular judgments. In Figure 1B, observer does not have access to information about angular subtense and must calculate angular size indirectly from objective size and depth, thereby lowering the precision of angular judgments relative to objective judgments.

Figure 2

Diagram of target used for size thresholds. Subjects judged incremental changes in the vertical distance separating the horizontal lines relative to an implicit reference (the mean of the set of test stimuli). For the objective size judgments in cm, target disparity was varied at random from trial-to-trial, and the angular subtense was scaled with target disparity to be consistent with the apparent physical distance of the target. For the angular size judgments in min arc, target disparity was varied at random from trial-to-trial, but the mean angular subtense did not vary with target disparity.

Figure 3

Diagram of experimental set-up. Left side of figure shows the apparent distances associated with the disparities used to measure angular and objective size thresholds. Right side shows the actual electronic stereoscope used to present stereoscopic display. See methods section for details.

Figure 4

Increment thresholds for vertical separation for three experimental conditions: Objective size judgments with random trial-to-trial changes in disparity; Angular size judgments with random trial-to-trial changes in disparity; Size judgments for target presented only in Fixation Plane. Target Duration 150 msec.

Figure 5

Ratio of Objective to Angular Size Thresholds as a function of mean angular size, equal to the angle subtended by the mean objective size at 1.5 meters. Target disparity varied at random from trial-to-trial. Data from subjects LW and SM, replotted from Figure 4, and is based on target duration of 150 msec. Data from subject WAM is based on target duration of 1000 msec.

Figure 6

The Points of Subjective Equality (stimulus value corresponding to 50% point on psychometric function) plotted as a function of target disparity for both angular and objective size judgments at two different sizes (5 and 30 arc min, or 0.22 and 1.3 cm). The oblique line in each graph shows the predicted change in angular subtense for objective size, the horizontal lines being the prediction for the angular size. Subjects are unable to compensate for depth at small scales. Target duration 150 msec

Figure 7

Ratio of Angular Size Thresholds, measured with random changes in disparity, to Size Thresholds measured only in the Fixation Plane. Target duration 150 msec The horizontal line shows mean ratio for these measurements. The random disparity thresholds for angular size are consistently higher than size thresholds measured in a single plane.

Figure 8

The P.S.E's plotted as a function of target disparity for both angular and objective size judgments (30 arc min or 1.3 cm) for two different durations (150 and 1500 msec). The longer duration improves the Weber fractions for both angular and objective sizes, shown in the boxes in the lower right corner of each graph. Increasing the target duration to permit a shift in convergence does not remove the difference between the single plane and multi-plane thresholds.

Figure 9

The Weber fractions for subject SM for 2 angular sizes and two different durations. At the longer duration, subject converged to the plane of the target. In the "One Plane" condition, the target was presented only in the fixation plane. In the "Two Planes" condition, the target was presented at random either in the fixation plane or with 30 min uncrossed disparity. In the "Nine Planes" condition, the target was presented at random in one of nine depth planes, covering $\pm$ 40 arc min range. For the small angular size (5 arc min), target disparity accounts for the loss of precision in both multi-plane conditions ("Two" or "Nine"); given time to converge, all conditions produce equally precise judgments. For the large angular sizes (30 arc min), target disparity has virtually no effect on the thresholds, but increasing the number of different depth planes degrades performance.

Figure 10

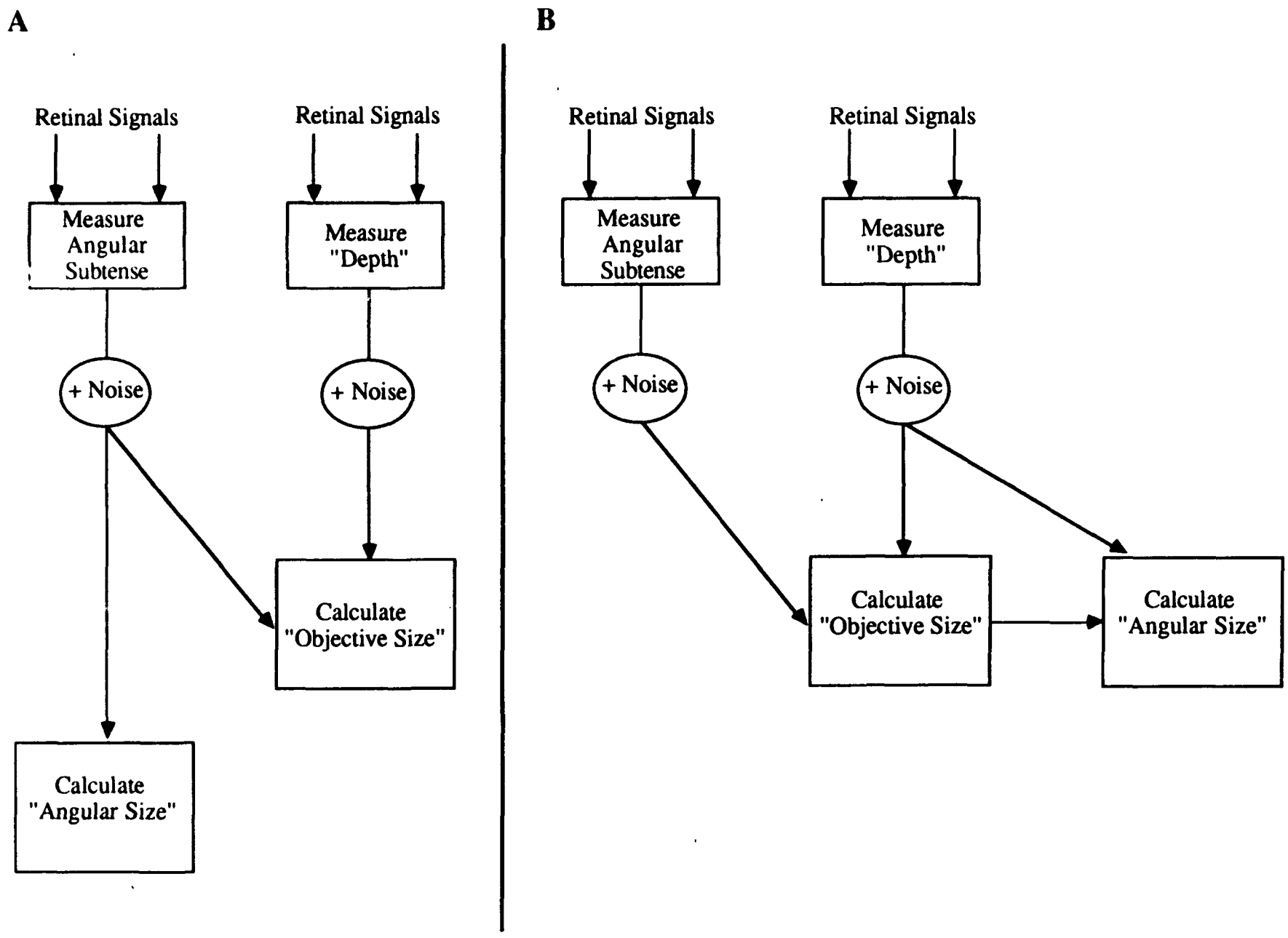
The P.S.E.'s for constancy and "anti-constancy" as a function of target disparity for one size (1.3 cm subtending 30 arc min at 1.5 meters). In the constancy condition, the angular subtense decreased with increasing depth consistent with the natural change associated with increasing physical distance. In the "anti-constancy" condition, the angular subtense increased with increasing depth completely inverting the natural relationship. Target duration 150 msec.

Figure 11

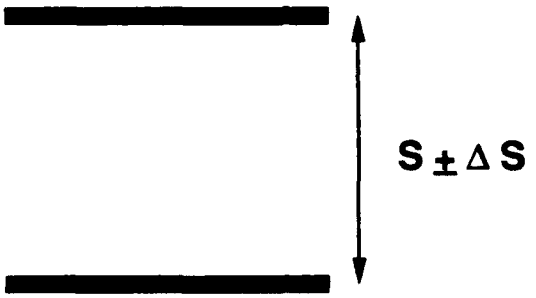
The Weber fractions for 30 min angular subtense, for the conditions in which target disparity changed from trial-to-trial to one of nine depth planes chosen at random: Angular size judgments; Objective size judgments ("Constancy"), and "Anti-Constancy" judgments.

Figure 12

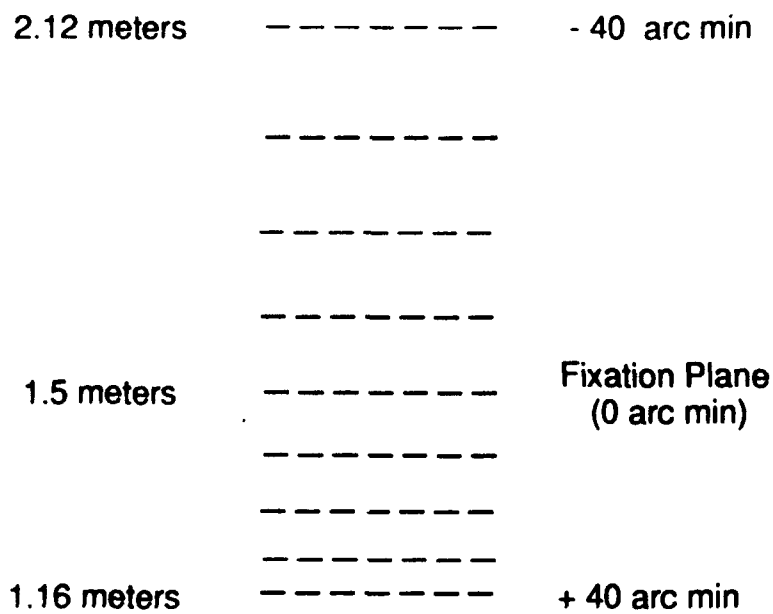
Flow Chart showing size processing at different scales; See Text.


A
Retinal Signals
Retinal Signals
Measure Angular Subtense
Measure "Depth"
+ Noise
+ Noise
Calculate "Objective Size"
Calculate "Angular Size"
B
Retinal Signals
Retinal Signals
Measure Angular Subtense
Measure "Depth"
+ Noise
+ Noise
Calculate "Objective Size"
Calculate "Angular Size"

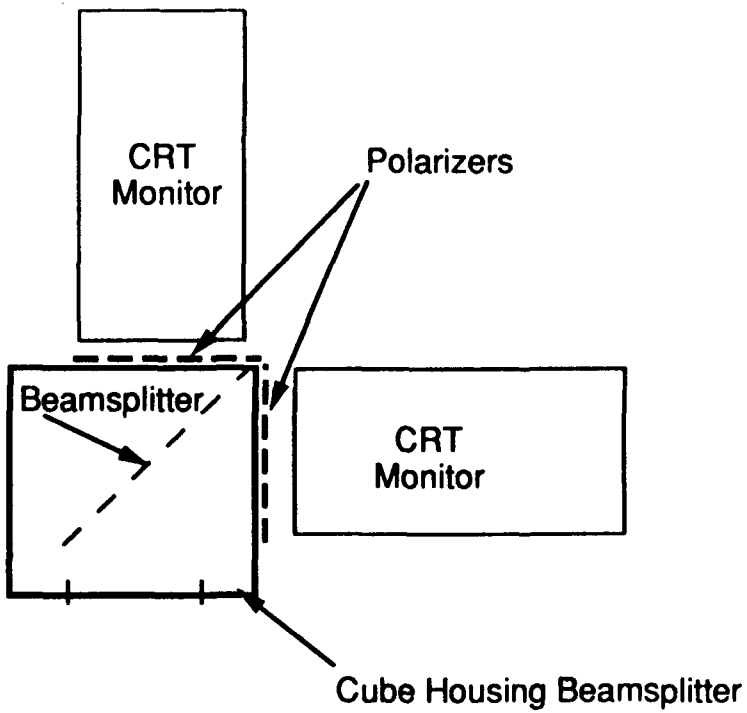
Test Target

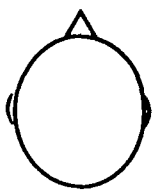


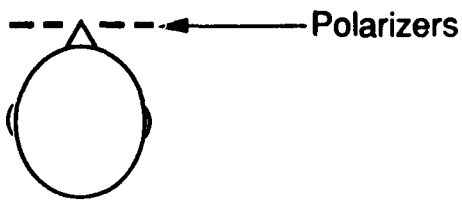
DISTANCES

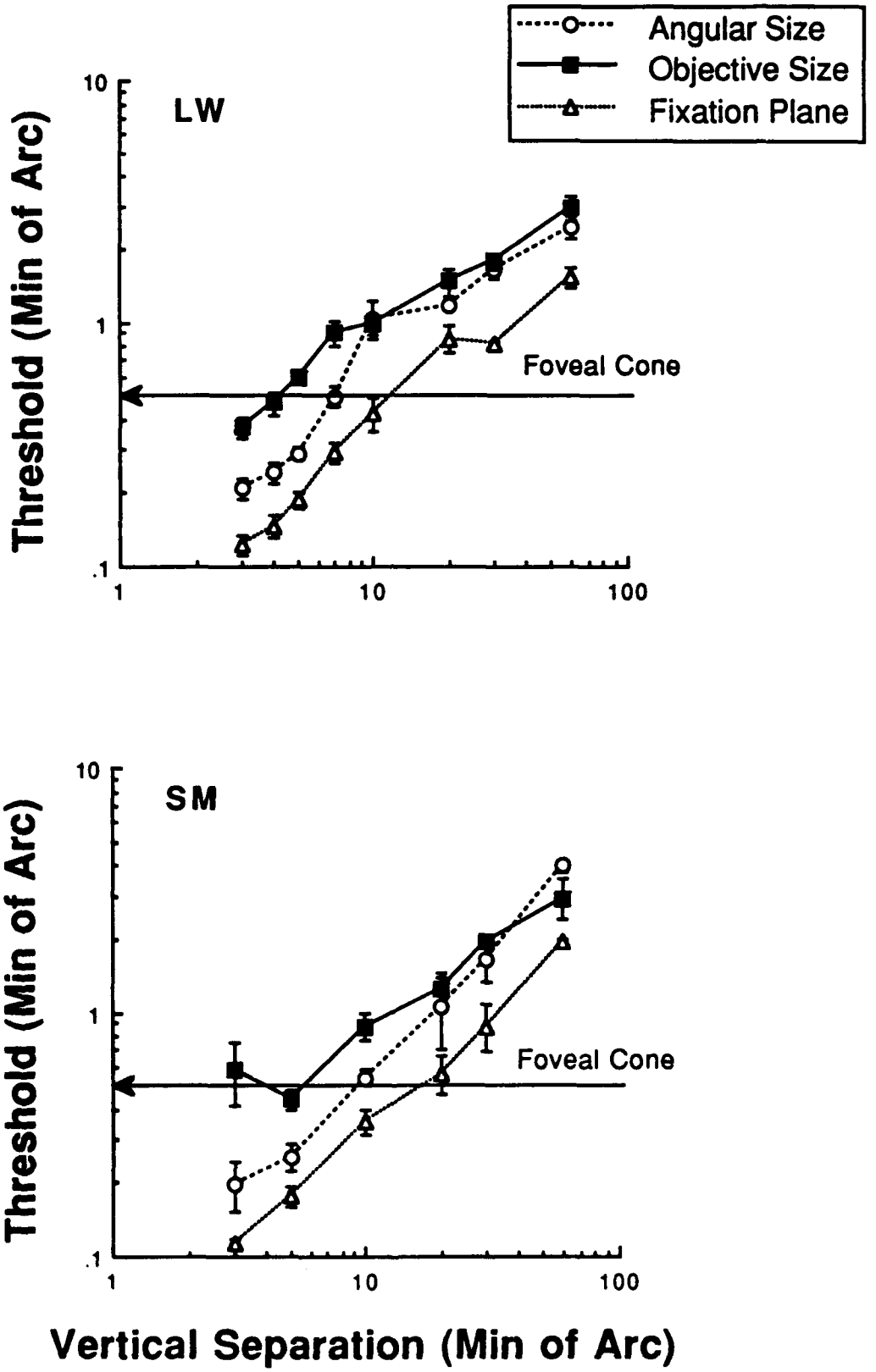


EXPERIMENTAL SET-UP

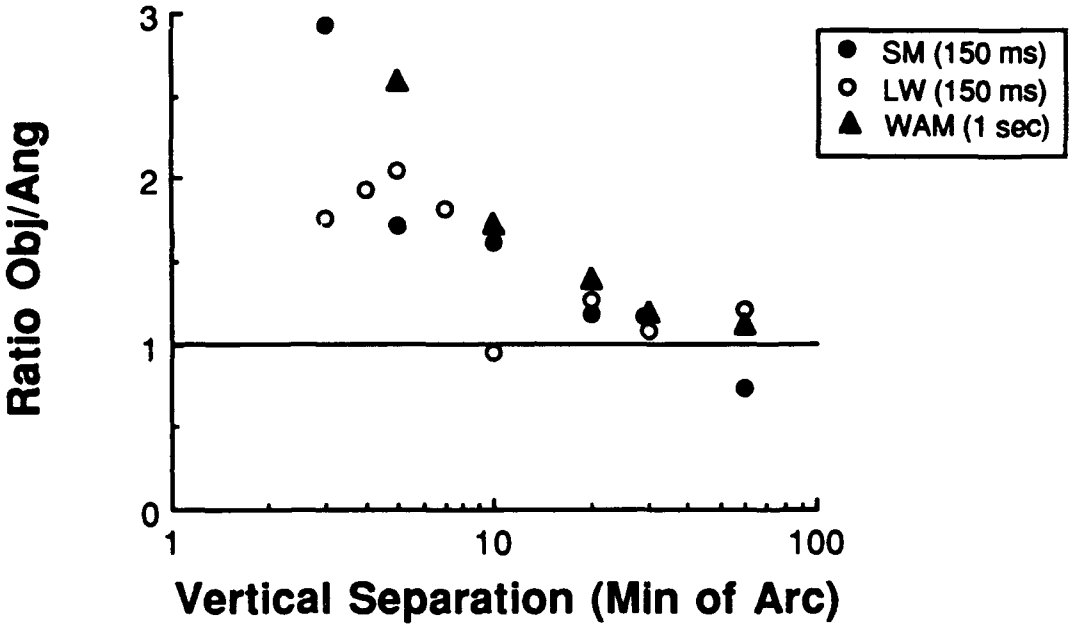


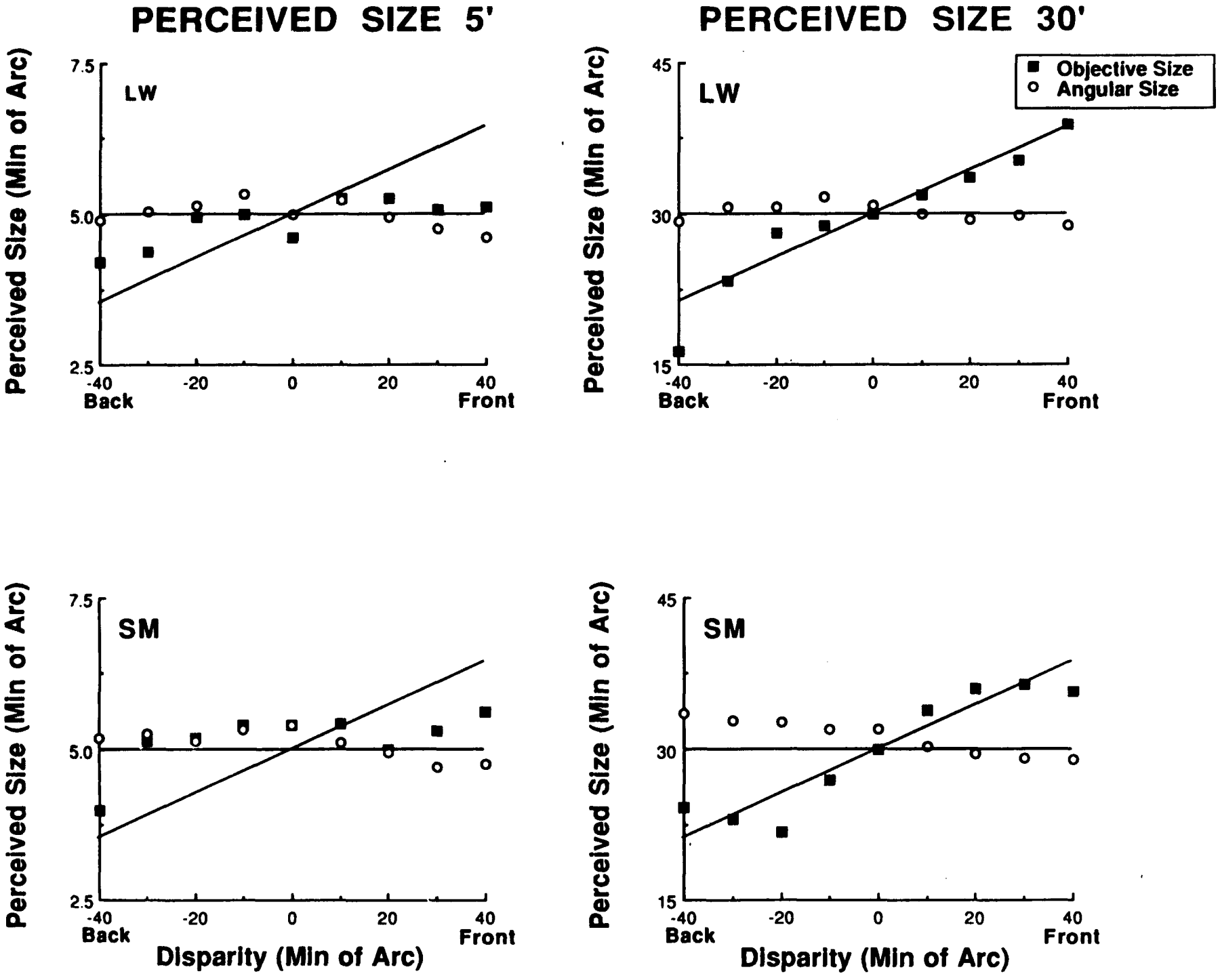




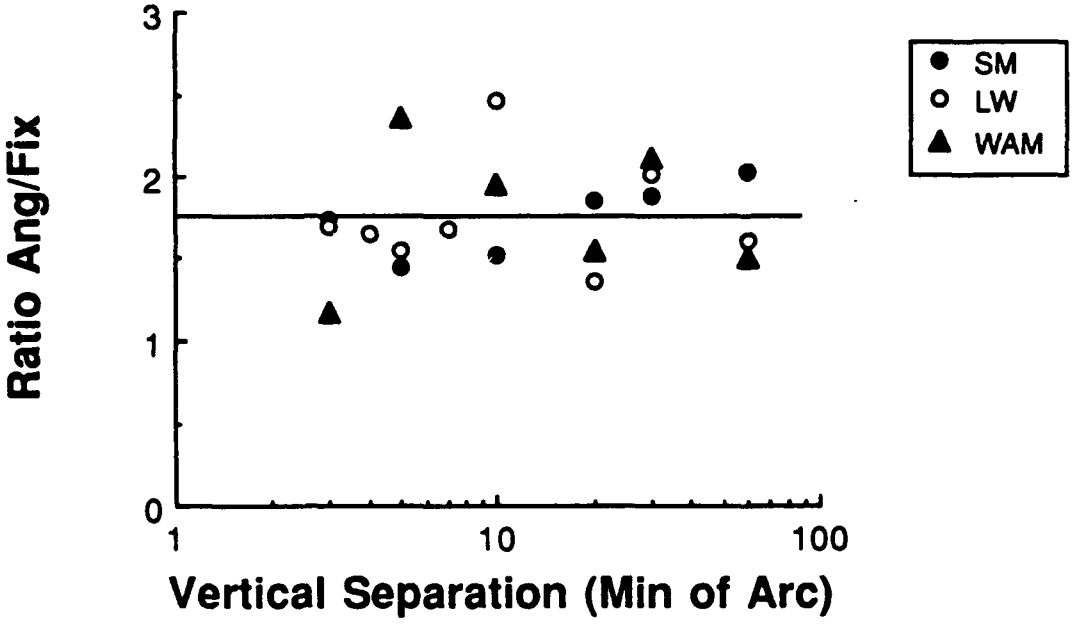

Angular Size
Objective Size
Fixation Plane
LW
SM
Threshold (Min of Arc)
10
1
.1
Foveal Cone
1
10
100
Vertical Separation (Min of Arc)

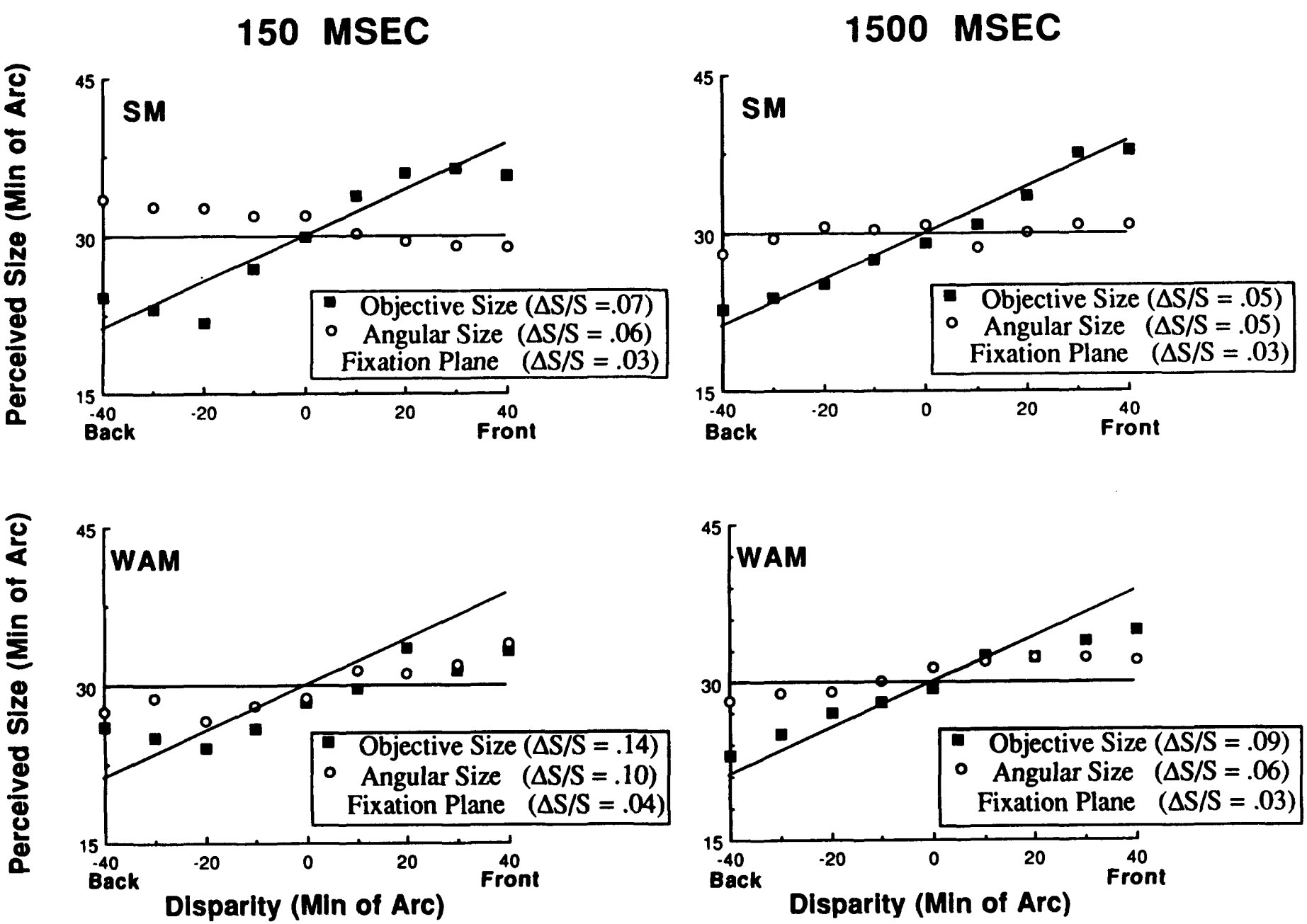
RATIO OF OBJECTIVE/ANGULAR THRESHOLDS




PERCEIVED SIZE 5'
PERCEIVED SIZE 30'
Perceived Size (Min of Arc)
LW
SM
Objective Size
Angular Size
7.5
5.0
2.5
45
30
15
-40
-20
0
20
40
Back
Front
Disparity (Min of Arc)

RATIO OF RANDOM DISPARITY ANGULAR THRESHOLDS TO FIXATION PLANE THRESHOLDS




150 MSEC
1500 MSEC
SM
WAM
Perceived Size (Min of Arc)
45
30
15
-40
-20
0
20
40
Back
Front
Objective Size (ΔS/S =.07)
Angular Size (ΔS/S = .06)
Fixation Plane (ΔS/S = .03)
Objective Size (ΔS/S = .05)
Angular Size (ΔS/S = .05)
Fixation Plane (ΔS/S = .03)
Objective Size (ΔS/S = .14)
Angular Size (ΔS/S = .10)
Fixation Plane (ΔS/S = .04)
Objective Size (ΔS/S = .09)
Angular Size (ΔS/S = .06)
Fixation Plane (ΔS/S = .03)
Disparity (Min of Arc)

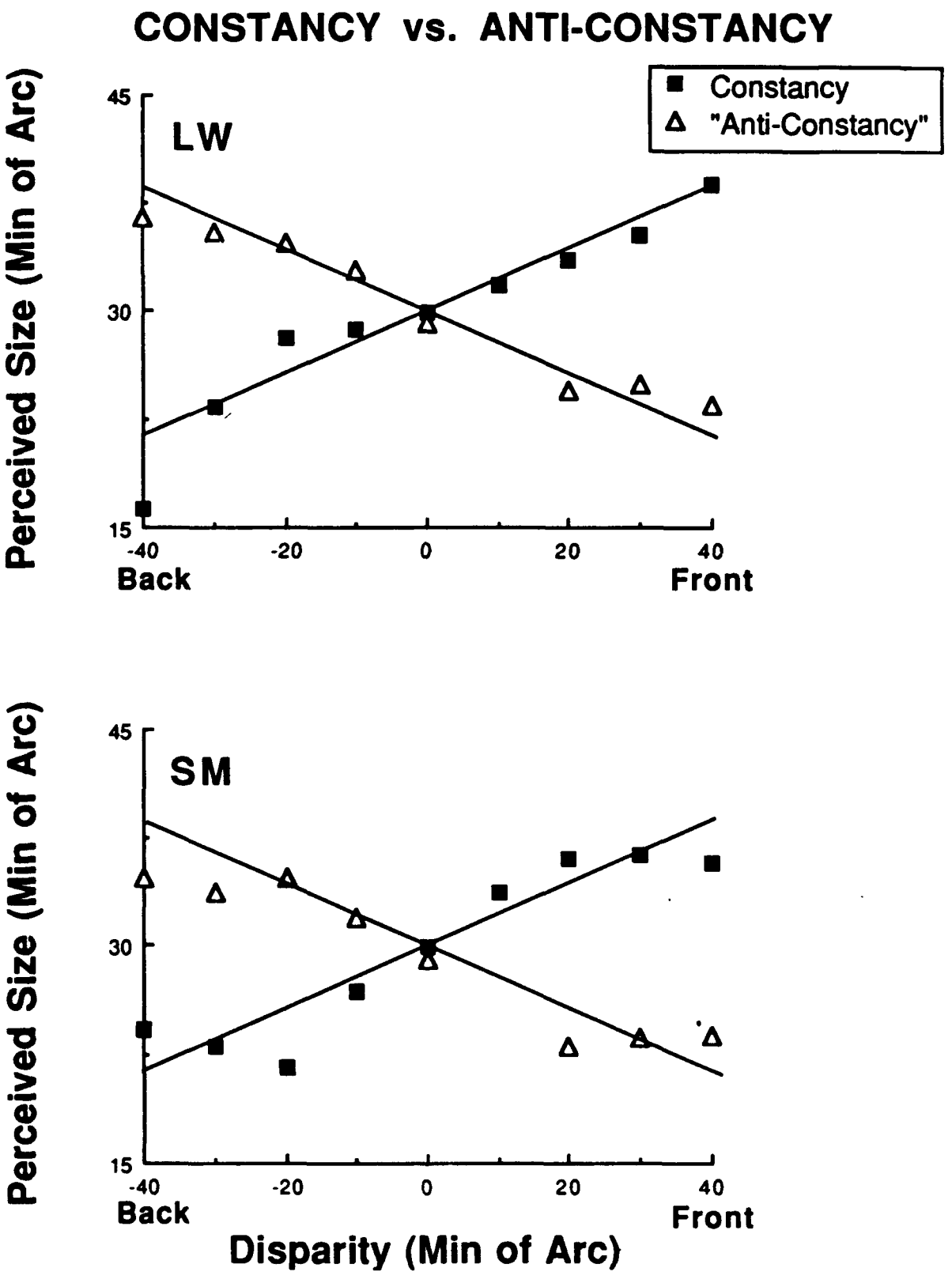
WEBER FRACTIONS FOR ANGULAR JUDGMENTS

5 MIN MEAN SEPARATION

Number of Planes	150 ms Duration	1500 ms Duration
One Plane	.041 (.002)	.031 (.003)
Two Planes	.056 (.005)	.039 (.004)
Nine Planes	.057 (.003)	.034 (.003)

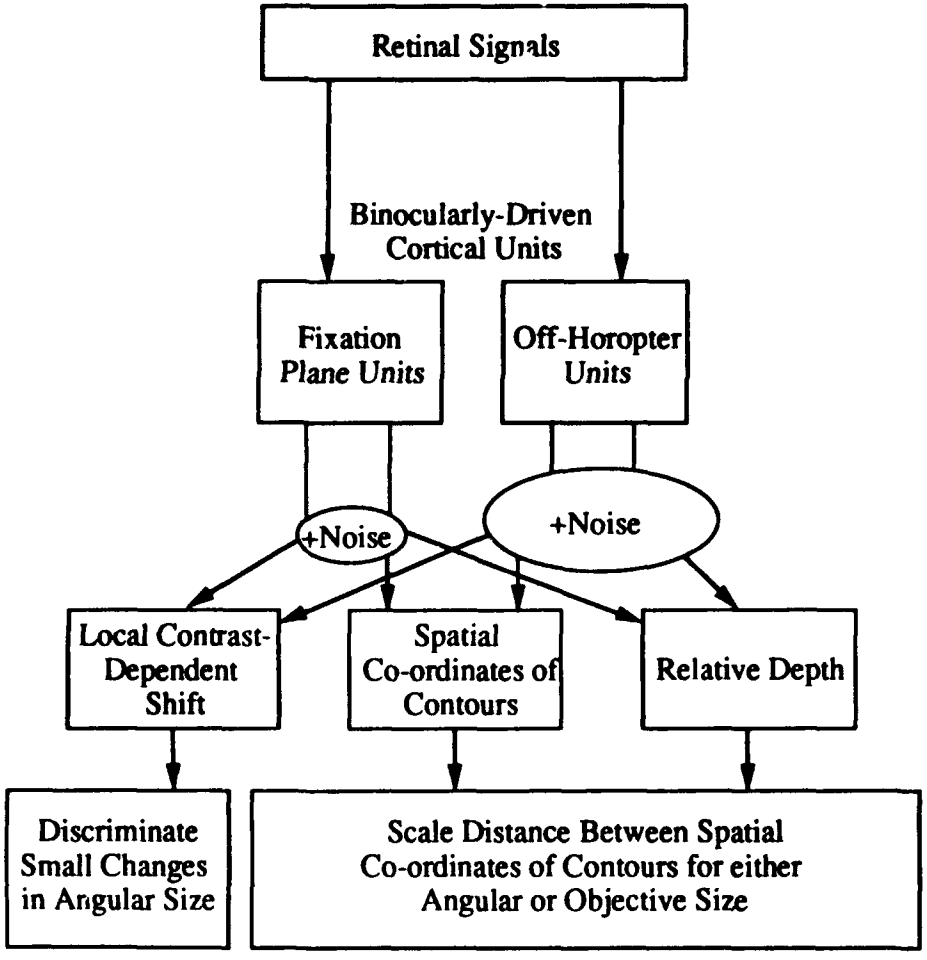
30 MIN MEAN SEPARATION

Number of Planes	150 ms Duration	1500 ms Duration
One Plane	.026 (.002)	.030 (.001)
Two Planes	.027 (.004)	.034 (.003)
Nine Planes	.052 (.003)	.043 (.002)

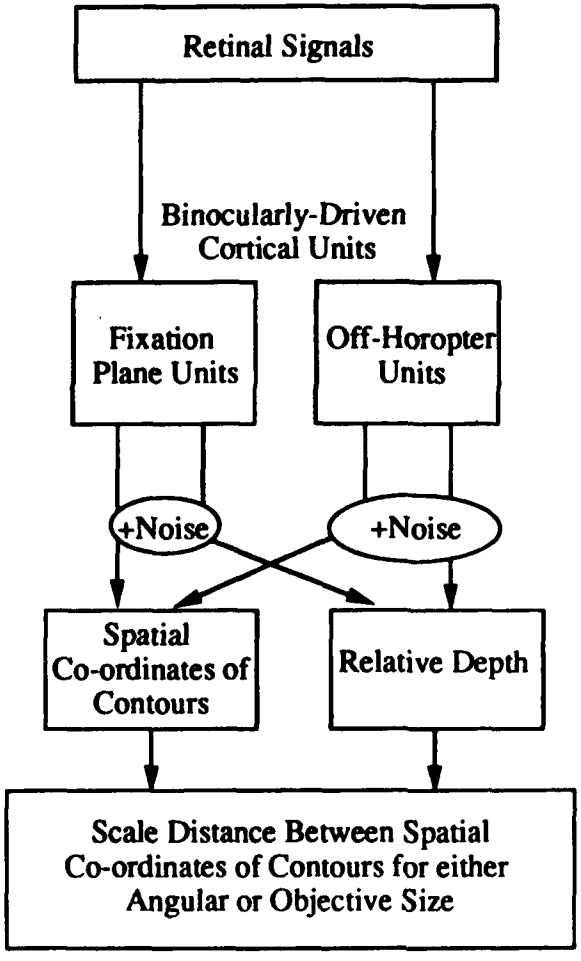

CONSTANCY vs. ANTI-CONSTANCY
Constancy
"Anti-Constancy"
LW
SM
Perceived Size (Min of Arc)
45
30
15
-40
-20
0
20
40
Back
Front
Disparity (Min of Arc)

	Many Planes Angular	Many Planes Objective	Many Planes "Anti-Constancy"
SPM 30' Separation	.059 (.004)	.065 (.004)	.08 (.006)
LW 30' Separation	.055 (.005)	.059 (.004)	.10 (.006)

FOR SMALL SIZES OR DISTANCES (< 10 MINUTES OF ARC)



FOR LARGE SIZES OR DISTANCES (> 10-20 MINUTES OF ARC)



The Human Visual System Averages Speed Information

Scott N. J. Watamaniuk[†‡] & Andrew Duchon[§]

Northwestern University, Department of Psychology, Evanston, IL 60208

Abstract - It has been known for many years that human observers are unable to detect modest accelerations and decelerations in moving visual stimuli (Gottsdanker, 1956). We find that human observers can integrate speeds over many dots, moving at different speeds, producing a global speed percept analogous to the global direction percept first reported by Williams and Sekuler (1984). We measured speed discrimination for random dot stimuli which contained many different speeds. Our results show that observers always base their discrimination on the mean speed of the stimulus; changes in other stimulus characteristics (e.g. mode) are not detected. Moreover, discrimination thresholds for the global mean speed derived from many different speeds are comparable to those obtained with stimuli in which all dots move at the same speed suggesting that the internal noise associated with the encoding of speed information is quite high.

Keywords: Integration Speed Discrimination Velocity Random Dots Global Motion Motion

Running Title: Humans Average Speed Information

Present addresses:

[†]The Smith-Kettlewell Eye Research Institute, 2232 Webster St, San Francisco, CA U.S.A 94115

[§]ATR Auditory and Visual Perception Research Laboratories, Kyoto 619-02, Japan.

[‡]To whom correspondence should be addressed.

Acknowledgments-We would like to thank Drs. Samuel Bowne, Mary Bravo, Suzanne McKee, Michael Morgan and Robert Sekuler for their helpful comments and suggestions. This research was supported by grants AFSOR-89-0243 and AFSOR-89-0035.

Introduction

The ability of the motion system to integrate information has been shown to be quite remarkable. Williams and Sekuler (1984) demonstrated that when humans are shown a stimulus containing many different local motion vectors, a unified global percept in the direction of the mean of the component directions may arise if the range of component directions is 180 deg or less. Moreover, the precision with which human observers can discriminate this global direction percept is very good — one to two degrees for distributions containing up to about 45 different directions (Watamaniuk, Sekuler and Williams, 1989). As well, Williams, Tweten and Sekuler (1991) found that observers could not distinguish between the percept produced by a uniform distribution of directions spanning 180 degs and that produced by a distribution containing only eight directions distributed evenly over the same 180 deg range.

Given that the visual system readily integrates direction information, a reasonable question is whether speed information is similarly integrated and whether the same constraints apply to both. There is evidence that the human visual system does integrate velocity information. Gottsdanker (1956) showed that humans are unaware of modest stimulus accelerations and this finding has been confirmed in many subsequent studies (Schmerler, 1976; Morgan, 1976; 1980; McKee and Nakayama, 1988; Bowne, McKee and Glaser, 1989; Snowden and Braddick, 1991). Gottsdanker speculated that such findings reflected an averaging or integration of preceding velocities.

The present experiments were designed to examine the nature of the integration of speed information and its precision. We chose to use random dot stimuli similar to those used by Watamaniuk *et al.* (1989), replacing the distribution of directions with a distribution of speeds with all dots moving upwards in the same direction. There are three goals for this paper: to determine 1) what characteristics of the statistical distribution of speeds oobservers use to perform speed discrimination, 2) the precision of speed discrimination with these random dot stimuli and 3) whether the human visual system integrates speed information across many dots.

We measured speed discrimination for stimuli containing a distribution of speeds, using a two-alternative forced-choice (2AFC) procedure, and independently varied the stimuli's mean and the modal speeds. In some conditions, the range of speeds spanned by a stimulus was held constant, while in others, the range of speeds covaried with the mean. The logic of this approach is that performance should change only when the relevant stimulus characteristic is changed. For example, if global speed discrimination depends only on the mean of the underlying distribution of speeds, then performance should improve as the difference between the mean speeds of the two stimuli gets larger; performance should be unaffected by changes in the mode.

Experiment 1: Speed Discrimination for Distributions of Speed

We wanted to test Gottstanker's idea that observers average speed information but with a stimulus in which the speed information is spatially distributed over many dots. This experiment was designed to determine how the visual system integrates speed information. Many stimulus conditions with different distribution characteristics were used in these experiments to determine if speed discrimination was affected by particular stimulus characteristics.

METHOD

Stimuli

Stimuli were dynamic random dot cinematograms. Each dot took a one-dimensional (1-D) random-walk in which each dot's displacement for each frame was chosen randomly from a predefined distribution of speeds and was independent of both its previous displacements and the displacements of other dots. Stimuli were created such that the underlying distribution of speeds was perfectly represented in each frame of the cinematogram. This was accomplished by having each element choose a speed from the distribution without replacement. Thus the stimulus contained every speed that appeared in the underlying distribution and at the appropriate frequency. Thus, the characteristics of each tested stimulus distribution were exactly specified. This sampling technique allowed for precise manipulation of the stimulus characteristics such as the mean

and modal speeds. All dots moved in the same direction, upwards.

In this experiment, the range of speeds in the distributions was constant for all stimulus conditions, spanning from 2.2 to 8.5 deg/sec. Keeping the range of speeds (or extreme speeds) constant in all trials prevented the observers from using the fastest or slowest speed as the basis of their speed judgments.

Specific Stimulus Conditions

Most of the distributions of speeds in this experiment had the same basic construction: a rectangular distribution and a triangular distribution atop a uniform background distribution. For all stimuli of this type, the uniform background distribution contained 64 different speeds (2.2-8.5 deg/sec in 0.1 deg/sec steps). Each speed of the background distribution was represented by either 1 or 2 dots, depending upon the stimulus condition, so that 64 or 128 dots moved with speeds selected from the background distribution. Additional dots were assigned to the speeds contained in the rectangular and triangular distributions with the maximum number of dots presented in each frame being 256. For each stimulus condition, we chose one speed to be the standard. Four comparison stimuli were constructed so that the mean or mode varied from the standard by about 5, 10, 15, and 20% although exact values varied slightly for different conditions.

We have grouped stimulus conditions based on the stimulus characteristics that changed between the standard and comparison, i.e. mean speed changed while mode was constant. The following are more detailed descriptions of each stimulus group.

a) Mean Changing with Mode Constant

The stimuli in this group were characterized by their mean speed changing from standard to comparison while their modal speed remained constant. Figure 1A shows an example of how the mean of these stimuli was increased, by shifting the rectangular distribution, while the mode, defined by the triangular distribution, was kept constant. Two different heights of triangular distributions

were used (7 and 14 dots) and each was presented in two forms: one with the rectangular distribution at the slow end and the triangular distribution at the fast end of the background distribution and the other with the rectangular and triangular distributions reversed. Density was 2.25 dots/deg^2 when the height of triangular distribution was 7 dots and 2.26 dots/deg^2 when the height of triangular distribution was 14 dots.

Figure 1. about here

---

b) *Mode Changing with Mean Constant*

This group comprised four stimulus conditions whose construction was identical to those previously described, but in these conditions the comparison's mode was changed by shifting the triangular distribution. Any change in the mean speed produced by the shifting of the triangular distribution was compensated for by shifting portions of the rectangular distribution. Thus mean speed was kept constant. Figure 1B snows shows an example of one such stimulus.

c) Mean and Mode Change Together

Stimulus conditions in this group were characterized by their mean and modal speeds covarying in the same direction — faster or slower. Two of these stimulus conditions had constructions identical to those previously described, a rectangular and triangular distribution atop a background distribution of speeds. Another two stimulus conditions had only a triangular distribution, of height 7 or 14, atop the usual background distribution. For all of these stimulus conditions, the mode was varied by shifting the triangular distribution which necessarily shifted the mean in the same direction. Density for the latter two conditions was 1.77 dots/deg^2 when the height of triangular distribution was 7 dots and 1.62 dots/deg^2 when the height of triangular distribution was 14 dots.

d) Mean and Mode Change Opposite to each other

The two stimulus conditions in this group were constructed identical to those previously described in section *a*, a rectangular and triangular distribution atop a background distribution. The mean and modal speeds were varied in opposite directions by shifting the triangular distribution one way and the rectangular distribution the other way. One condition had mean speed increasing and modal speed decreasing while the other had the reverse. Figure 1C shows an example of one of these stimuli.

No feedback was provided with these two conditions to avoid biasing responses. Providing feedback based on either the mean or modal speed necessarily forces the observers to choose a particular response strategy to avoid making an undue number of mistakes. For example, if the feedback was based on mean speed but the modal speed was the potent cue, the observers would begin to choose the stimulus that moved slower, rather than the one that moved faster, to avoid making mistakes.

e) Mean and Range of Speeds Change Together

The stimuli in this condition were different from those described above because they were uniform distributions, of height 4, whose standard spanned the same range of speeds as the previous stimulus conditions producing a mean speed of 5.34 deg/sec. To change the mean speed, the entire distribution was shifted resulting in the range of speeds being shifted by the same amount and in the same direction as the mean speed. These stimuli had densities of 2.56 dots/deg^2.

Apparatus

Stimuli were displayed on an x-y cathode ray tube display (CRT) with a fast, P4, phosphor. A circular mask with a diameter of 9 degrees of visual angle was placed over the 10x10 deg CRT screen. Each dot subtended about 6 min. arc and had a luminance of about 0.27 cd/m^2.[1] The background and veiling

[1]This value was obtained by plotting a matrix of non-overlapping dots (center-to-center spacing was 0.06 deg) at the same frame rate as used in the experiments. The luminance of this matrix was then measured with a Minolta luminance meter. The decay rate of the phosphor

luminances were 0.03 and 0.07 cd/m^2, respectively.

The observer, viewed the CRT from a distance of 57 cm. The height of the CRT was set so that the center of the aperture was at approximately eye level. Observers fixated on a spot located at the center of the aperture. Push buttons connected to a computer initiated each trial and signalled observer responses. All experiments took place in a darkened room and before testing, observers were allowed five minutes for their eyes to adapt.

Procedure

Stimuli were presented within a 2AFC paradigm. In each trial, two cinematograms were displayed successively with a blank ISI of about 300 msec. For each stimulus condition, one of the cinematograms was chosen to be the standard. This stimulus appeared in every trial. The other cinematogram, called the comparison, differed from the standard in its mean speed, modal speed or both. The order in which the standard and the comparison were presented was randomized from trial to trial. Although the stimuli were shown at a constant frame rate of 20 Hz, the duration of the cinematograms was varied randomly, from 250 to 450 msec, so that the observers could not base their decision on the distance travelled by the elements. As well, the two cinematograms presented within any trial were forced to have different durations.

The observers task was to choose the cinematogram that had the faster speed. Each observer completed 200 trials for each standard/comparison pairing for every stimulus condition. The order in which the different stimuli were tested was randomized. Ten practice trials were presented at the beginning of each set of five 100-trial blocks and each experimental session comprised ten blocks. For those conditions in which only the mean speed changed or other characteristics of the stimulus changed together with the mean, feedback based on the mean speed of the distributions was provided. For those conditions in which only the mode changed, feedback based on the modal speed was given. No

---

was such that the luminance of the dots decreased to 1% of their peak value after about 20 msec. Since each frame was approximately 50 msec, the luminance of each dot was essentially zero when its next position was plotted.

feedback was given when the mean and modal speeds were shifted in opposite directions.

Performance, in percent correct, was obtained for each standard-comparison pairing in each stimulus condition. A Weibull function was fit to each set of data by a maximum likelihood technique which directly searched through a list of many, closely spaced, threshold values. Thresholds, corresponding to the Weber fractions ($\Delta V/V$) necessary to produce performance corresponding to d'=1.0, were evaluated from the fitted curves. The standard error of each threshold was estimated as in Quest (Watson and Pelli, 1983).

Observers

One of the authors (AD) and two university students (DB & WS) served as observers. All observers had normal or corrected to normal vision. Observers completed 2000-3000 training trials to accustom them to the task and were tested over a period of several weeks.

Results

Average Weber fractions for all but the 'mode-changing & mean-constant' stimulus group appear in Figure 2. There is one major point about these data that we want to make; observers' discrimination was <u>always</u> based on the change in mean speed. Performance did not change significantly with stimulus condition (confirmed with an ANOVA, $F_{3,35} = 1.69$: $p=0.1868$) so that whether the mean speed changed alone or the mean and mode changed in the opposite or the same direction, performance did not change.

Figure 2. about here

When only the modal speed changed and the mean speed was held constant, observers performed at chance levels even for mode changes as large as 20%. Figure 3 plots performance in percent correct, averaged over observers, as a function of the magnitude of change in modal speed.

Figure 3. about here

These results strongly suggest that the human visual system can extract the mean speed of a stimulus composed of many different speeds and use that information as the basis for discrimination. This result is consistent with past findings (Gottsdanker, 1956; Schmerler, 1976; Bowne et al., 1989; Morgan, 1976, 1980) and unequivocally shows that the human visual system averages speed information to derive the mean. Accurate information about the modal speed is not available and shifting the entire distribution, as when the mean & range changed together, does not seem to provide much additional information.

Given that observers can judge the mean speed of these speed distributions so well, it is interesting to note that when viewing the stimuli, one does not perceive all the dots moving at a single common speed — dots are seen moving at different speeds (generally one sees a group of dots moving fast and a group moving slow). This observation is curious because it seems to be at odds with our discrimination results.

Experiment 2: Effect of Stimulus Variance on Speed Discrimination

The results of the previous experiments showed that speed discrimination was essentially unaffected by characteristics of the stimulus other than the mean speed: performance was the same across all stimulus conditions. This is not too mysterious as the previous stimuli were produced in such a way that the distributions of speeds were always specified exactly within each frame of each presentation. If the variability of the visual system's encoding of speed was independent of the actual speed, within the range of speeds used here, then one would not expect a change in discrimination because there is no variability in the stimulus. We decided to test how much variability in the visual system was associated with the encoding of speed by measuring how sensitive speed discrimination was to stimulus variability. The idea is that speed discrimination should change only when the stimulus variability exceeds the variability inherent

in the visual system.

Stimuli

Speed discrimination was measured for five Gaussian-shaped[2] speed distributions with standard deviations (SD) of 0.0[3], 0.43, 0.85, 1.27 and 1.70 deg/sec. In terms of the range of speeds spanned by these stimuli, a stimulus with a mean speed of 7.6 deg/sec with an SD=0.43 would have speeds ranging from 6.4 deg/sec to 8.8 deg/sec, while one with the same mean speed but with an SD=1.70 would contain speeds from 3.7 deg/sec to 11.5 deg/sec. These stimuli had densities of 2.56 dots/deg^2.

Like the uniform speed distribution described previously, the mean speeds of these distributions were varied by shifting the entire distribution. However, unlike the previous stimuli, speeds were randomly assigned to the stimulus dots with replacement. Sampling with replacement resulted in a distribution of speeds for any one frame that was a random sample of the underlying speed distribution. Thus Gaussian distributions of speeds with larger SDs produced stimuli with larger variance.

All of the standard stimuli for the Gaussian conditions had mean speeds of 7.6 deg/sec. For each of the five Gaussian conditions, there were four comparison stimuli that were faster and four comparison stimuli that were slower than the standard by 5, 10 , 15 and 20%. Each observer completed 200 trials for each standard/comparison pairing for each of the five Gaussian conditions. The two cinematograms presented in any trial always had the same SD.

The apparatus and other procedures were identical to those described in Experiment 1.

Observers

Two observers from the previous experiment (AD & WS) served as

[2]Because of the discrete nature of the display, it was not possible to present a continuum of speeds. We approximated a Gaussian distribution of speeds by sampling at 0.1 deg/sec intervals.

[3]The stimulus with an SD=0.0 deg/sec refers to one in which all elements moved at the same speed.

observers for this experiment.

Results

Performance, in percent correct, was obtained for each standard-comparison pairing in each stimulus condition. Because performance was symmetric about the standard speed, the data were combined and analysed as a function of the absolute change in mean speed. Therefore, there were 400 trials for each data point. A Weibull function was fit to each data set and thresholds, corresponding to the Weber fraction ($\Delta V/V$) necessary to produce performance corresponding to d'=1.0, were evaluated from the fitted curves.

Figure 4 plots the thresholds for each Gaussian condition, averaged over observers, as a function of the stimulus SD. It is obvious that discrimination thresholds were the same across all Gaussian conditions (confirmed by ANOVA, $F_{4,5} = 0.435$, $p=0.78$). Performance was unaffected by increasing the variability of the stimulus; performance was equally good for a stimulus containing one speed (SD=0.0, no variability) as one spanning an 8 deg/sec wide range of speeds (SD=1.70). Thus increasing stimulus variance or noise over this range did not affect speed discrimination. This result suggests that the internal noise associated with the encoding of speed is quite large - at least larger than the variance of the present stimuli.

Figure 4. about here

Do Humans Average Speed Over Many Dots or Just One?

The previous experiments showed that when presented with stimuli containing many dots moving at different speeds, humans are able to discriminate those stimuli based on the mean speeds. How many dots is this average speed based on? Do humans average speeds over only a single dot or do they average the speeds of many dots together?

Since the previous stimuli were created using the 1-D random-walk algorithm, which randomly assigns each dot a new speed from the distribution

each frame, the average speed of any one dot may have been sufficiently representative of the average speed of the entire display to allow good performance[4]. Indeed past research has shown that human observers are unaware of modest changes in the speed of single moving objects (Gottsdanker, 1956; Schmerler, 1976; Bowne et al., 1989) or a field of dots in which all dots move at the same speed (Snowden & Braddick, 1991). Thus it is quite possible that our observers may have used this "single-dot" strategy to make their judgements. We decided to test whether our observers were averaging speed over one or many dots in two ways: 1) by comparing our human data to that from a computer simulation using the single-dot strategy and 2) by conducting an experiment in which the single-dot strategy would fail.

Computer Simulation

In our computer simulation of speed discrimination, we replicated the same stimulus conditions as that experienced by our human observers in Experiment 1. To simplify our calculations, the computer simulation was only run for a stimulus duration of seven frames, which was the mean duration of the stimuli shown to our human observers. The computer simulation evaluated the probability of making a correct judgement knowing only the average speed of one randomly chosen dot in each of the two stimulus intervals presented within a trial. The simulation was run for the same number of trials per condition as obtained with our human observers for 13 stimulus conditions from the first experiment (excluding the mode change and mean constant conditions). An ANOVA showed that there was a significant difference between the human and computer performance ($F_{1,50} = 6.11$; $p=0.017$) with human observers having lower thresholds (better performance). This suggests that performance as good as that displayed by our observers is unlikely to arise from speed judgments based on only the average speed of a single dot; the speeds of several dots must be averaged together.

---

[4]We thank Professor M. J. Morgan for his insightful comments on this topic.

Experiment 3: Averaging Speeds Over Many Dots

The comparison with the computer simulation showed that our human observers performed better than that predicted by taking the average speed of a single dot in each stimulus interval. Perhaps observers used two or three dots to achieve this precision. How many dots would they need? We chose not to answer this question directly by measuring speed discrimination for stimuli with different numbers of dots because there undoubtedly would be a large dependence on the spatial arrangement of the dots. Rather, we chose to create a stimulus that required averaging speed over many dots in order to do the task.

Stimuli

Stimuli were dynamic random dot cinematograms, as in Experiment 1, but dots were displaced each frame using one of two types of movement algorithms. One was the 1-D random-walk algorithm described above. The other is referred to as the fixed-trajectory algorithm. In this algorithm, once a dot has been assigned a speed from the predefined distribution, it continues to move at that speed for the entire duration of the stimulus presentation. As in Experiment 1, stimuli were created such that the underlying distribution of speeds was perfectly represented in each frame of the cinematogram and all dots moved upwards. It is important to note that if the same underlying distribution of speeds was used, the distribution of displacements taken by the dots presented within any one frame would be identical regardless of which movement algorithm was employed. The movement algorithm only determines how the displacements are redistributed after each frame. Figure 5 shows a schematic representation each of the two movement algorithms.

Figure 5. about here

Speed discrimination was measured for two stimulus conditions taken from Experiment 1: one in which only the mean changed while the mode remained constant and another in which the mean and mode changed together. Discrimination for both conditions was measured separately for both movement

algorithms so that performance could be compared between the two. Observers completed 100 trials for each standard/comparison pairing for each stimulus condition. The apparatus and other procedural conditions were identical to that in Experiment 1.

Observers

One author (SW) and two other experienced psychophysical observers (SM &MB) provided data for this experiment.

Results

Performance, in percent correct, is plotted as a function of the percent change in mean speed for both stimulus conditions in Figure 6. Open symbols present the average of our three human observers while the closed symbols present the computer simulation data. Plus and minus one standard error are plotted on each point. There are three important points to be made: 1) human performance is always superior to the computer's single-dot strategy, 2) the fixed-trajectory movement algorithm dramatically reduces the computer's performance but 3) the 1-D random-walk and fixed-trajectory stimulus algorithms produce nearly identical discrimination performance in human observers: performance is unaffected by the way speeds are assigned to individual dots from frame-to-frame. These data suggest that observers can and do average speed information over many dots regardless of how the stimuli are generated.

Figure 6. about here

One may argue that the observers used a different strategy for each of the two movement algorithms: a single-dot strategy for the 1-D random-walk and averaging over many dots for the fixed-trajectory conditions. This seems unreasonable since the two algorithms produce virtually identical performance. However, to dispel this argument, we measured speed discrimination for the same two stimulus conditions as before but used the 1-D random-walk algorithm for one of the stimuli within a trial and the fixed-trajectory algorithm for the other.

We call this the mixed-algorithm condition. The same three observers completed 100 trials for each standard/comparison pairing. Figure 7 plots the average performance of our observers as a function of the percent change in mean speed. Data for the 1-D random-walk and the fixed-trajectory algorithms have been replotted from Figure 6 for comparison. It is obvious that the data obtained using the 1-D random-walk, fixed-trajectory and mixed-algorithm are not significantly different from each other (confirmed by ANOVA - Mean & Mode Change Together condition: $F_{2,42}=0.297$, $p=0.744$; Mean Change & Mode Constant condition: $F_{2,42}=0.067$, $p=0.935$). This provides strong evidence that our observers averaged together the speeds of many dots and responded to the globally-defined mean speed of each stimulus regardless of the movement algorithm used. These results are analogous to those reported for global direction discrimination by Watamaniuk *et al.* (1989). It is of interest to note that when performing the speed discrimination task, our observers were unable to identify which movement algorithm was being used even in the mixed-algorithm condition.

Precision of Speed Discrimination

We have established that human observers can integrate speed information and discriminate stimuli based on a globally-defined speed[5]. We now look at the precision of that integration. Previous studies have shown that human observers seem to have a remarkable ability to discriminate one speed from another. For example, McKee (1980) found that people can discriminate less than a 5% change in the speed of a moving bar (d'=0.67). Using random dot displays that presented hundreds of moving dots, De Bruyn and Orban (1988) obtained thresholds of 7-12% (d'=1.4) while Snowden and Braddick (1991) obtained thresholds of about 6% (d'=1.14). Such discrimination appears to be

[5]With our stimuli, it is not the case that all dots are perceived as moving at the same speed. Although the dots were seen flowing as a group, observers were aware that some of the dots moved at different speeds than others. This global speed percept is completely analogous to the global direction percept reported by Williams & Sekuler (1984) in which a field of dots seemed to flow *en masse* in a single direction although observers were aware that not all of the dots were moving in the same direction.

quite extraordinary, but these studies used redundant, noise-free stimuli. For example, in De Bruyn and Orban's stimuli as well as Snowden and Braddick's, all elements moved at the same velocity. We can compare the present data to those of other researchers by putting all of the thresholds to the same d' level. Choosing d'=0.67, we find that average thresholds from our experiments ranged from 5-10% while De Bruyn and Orban's range from about 3.5-6% and Snowden and Braddick's are about 4%. Our thresholds are comparable though somewhat higher. The slight elevation in the present thresholds may be due to the slow speeds present in our stimuli. Other researchers (McKee, 1980; De Bruyn and Orban, 1989; Snowden & Braddick, 1991) have found that differential speed thresholds can increase significantly for speeds slower than 4 deg/sec suggesting that slow speeds may not be encoded as precisely as faster speeds. This hypothesis is supported by our data — the thresholds obtained with the Gaussian speed distributions, in which the slowest speeds ranged from 3.7 deg/sec (SD=1.7) to 7.6 deg/sec (SD=0.0), are lower than those obtained with the other stimuli. Our thresholds, although a little higher than others, are still in good agreement with those of McKee, De Bruyn and Orban, and Snowden and Braddick.

Discussion

In the introduction, we asked 1) what characteristic of the statistical distribution of speeds do observers use to perform speed discrimination, 2) what is the precision of speed discrimination with these random dot stimuli and 3) can the human visual system integrate speed information across many dots. Our results show that when presented with a stimulus containing many different spatially-interspersed speeds, the visual system averages the speeds together. However, although the visual system evaluates the mean speed, a summary statistic, specific information about the underlying distribution is not known: observers cannot discriminate changes in the modal speed. Moreover, the averaging or integration is performed over many dots, not just one, generating a global speed percept. And finally, observers can discriminate the average speed

of a stimulus comprising many speeds with the same precision as when the stimulus contains only one speed.

But why does the visual system evaluate the mean speed? One logical reason for evaluating the mean speed is to reduce the noise in the signal. For example, in our random dot stimuli there is the possibility that elements, from one frame to the next, are mismatched. According to the nearest-neighbor rule of feature matching, ir one feature moves from location A to A', and another feature moves from location B to B', a mismatch will occur if the distance between A and B' is less than the distance between A and A' (Williams and Sekuler, 1984). Assuming this rule, the probability that such mismatches would occur in the present stimuli is between 0.037 (for the smallest step size) and 0.433 (for the largest step size). However, such mismatches would not affect the average velocity of the entire pattern of dots, if we assume the uniqueness constraint: that every element in one frame is matched with only a single element in the next frame. This can be proven algebraically as shown in Equation 1.

$$V_{AVG} = \frac{\langle \Delta X \rangle}{\Delta t} = \frac{\left(\frac{\sum_i (X'_i - X_i)}{N} \right)}{\Delta t} = \frac{\frac{\sum_i X'_i}{N} - \frac{\sum_i X_i}{N}}{\Delta t} = \frac{\langle X' \rangle - \langle X \rangle}{\Delta t} \qquad (1)$$

In this equation, V_{AVG} is the average speed of the stimulus, ΔX is the average displacement of the elements, Δt is the time over which the displacement takes place, X_i is the i^{th} element's initial position, X'_i is the i^{th} element's position after its displacement, and N is the number of displaced elements. Equation 1 shows that regardless of how elements are matched from frame-to-frame, if the uniqueness constraint is adhered to, the average velocity is determined entirely by the mean displacement. In contrast, it is possible that through mismatching, the modal or extreme speeds could be misperceived, thus these sources of information are necessarily noisier than the mean. Therefore the problem of mismatches is remedied only by calculating the mean speed. By this reasoning, the most reliable piece of information that can be extracted when presented with a distribution of speeds is the mean speed.

Perceiving a global mean speed when viewing a distribution of speeds is analogous to perceiving a global mean direction when viewing a distribution of directions. In fact, a closer examination shows that many of the details of speed integration are also seen in direction integration. First, our observers discriminated the global speed of stimuli containing many different speeds as precisely as ones containing only a single speed (see Figure 4). Similarly, Watamaniuk *et al.* (1989) showed that direction discrimination was equally good if the underlying direction distribution spanned one or 30 deg. Secondly, our data show that a uniform distribution of speeds (Mean & Range Change Together condition in Figure 2) produced slightly poorer performance than a Gaussian distribution spanning a similar range of speeds (SDs of 1.27 and 1.7 deg/sec in Figure 4). One possible reason for this is that because the Gaussian distributions have more energy at the mean speed than the uniform distribution, they produce a stronger mean-speed signal and subsequently better performance. This was also found by Watamaniuk *et al.* (1989) for global direction discrimination. And finally, we have shown that whether individual dots changed their speed randomly each frame (1-D random-walk) or moved at only a single speed for the duration of the presentation (fixed-trajectory), discrimination performance remained unchanged. This too was shown for direction discrimination by Watamaniuk *et al.* (1989).

The marked similarities between the present results and those for global direction discrimination suggest that the mechanism underlying both global processes are very similar. One type of model that has been successful in accounting for global direction performance has been a line-element model (Watamaniuk *et al.*, 1989; Williams *et al.*, 1991). Their line-element model was composed of a small number of bandlimited mechanisms, each sensitive to a particular range of directions. This model not only accounted for all of the details of discrimination performance described above but it also correctly predicted motion metamers -- the number of discretely sampled directions needed to create a stimulus that is indistinguishable from a stimulus comprised of a large range of continuously sampled directions. The line-element model developed by Williams and colleagues seems to capture well the perceptual qualities of global

direction.

The striking similarities between global speed and direction discrimination suggests that the same type of line-element model, with the bandlimited mechanisms selectively sensitive to a particular range of speeds rather than directions, would be a good candidate for describing global speed discrimination. In addition to the similarities described above, the result that observers can discriminate stimuli based on a change in mean speed but not modal speed, which was not explicitly tested by Williams and Sekuler (1985) or Watamaniuk *et al.* (1989) for global direction discrimination, is also consistent with a line-element model. For the asymmetrical speed distributions used in our experiments, being able to evaluate the mean speed requires that the visual system not only measure the many different speeds present but it must also know the frequency of occurrence for each speed -- evaluating the mean as if each different speed was represented equally (as in a uniform distribution) would result in an incorrect estimate. The same information about frequency of occurrence is necessary for evaluating the modal speed. However, when asked to discriminate global speed, observers can only discriminate changes in the mean speed. It seems that when evaluating global speed, accurate information about only the mean is available: the local speeds and their frequency (defining the underlying speed distribution) are obscured. Consequently changes in the modal speed are not discriminable. This is consistent with a line-element model because the composition of the underlying stimulus distribution is automatically accounted for by the way the model's mechanisms' responses are calculated. An individual mechanism's response is calculated by first multiplying its sensitivity to each speed component by the frequency of occurrence of that component in the stimulus and then summing those products. Once the summing is done and the mechanism's response is compressed to a single number, the individual components contributing to that response are obscured. This is consistent with our data and the fact that observers were unaware that different movement algorithms were being used. We hope to develop this line-element model of global speed discrimination in the future.

Finally, we would like to contrast our findings on global speed

discrimination to some of the ideas developed in the structure from motion (SFM) literature. We make this comparison because the stimuli used in SFM studies are very similar to those used in the present experiments, but the observer's task is different.

Several SFM studies have used random dot stimuli in which the dots moved as if they were on a transparent rotating cylinder. In these experiments, it is suggested that relative velocity information is used to extract a 3-D structure (i.e. Siegel & Andersen, 1988; Treue, Husain & Andersen, 1991). Interestingly, the stimuli need not be true parallel projections of points on the surface of transparent objects to produce a 3-D percept. Treue *et al.* (1991) reported that when velocities were randomized so that each dot no longer followed a trajectory corresponding to a veridical parallel projection of the surface of a rotating cylinder, some observers saw a transparent rotating cylinder that had dots distributed throughout. Williams and Phillips (1986) found that a similar percept resulted when viewing random dot cinematograms containing a distribution of directions rather than speeds. In their stimuli, all dots moved at the same speed but randomly chose their direction of movement from a distribution of directions.

Similarly, the present stimuli could have been interpreted by the visual system as 3-D objects. There are at least two rigid 3-D-object percepts that could produce a velocity field like that of our stimuli: 1) a rotating cylinder filled with dots, transparent to about the center, viewed through an aperture smaller than the cylinder's diameter or 2) many transparent sheets of random dots sliding over each other. But when observers concentrated on performing our speed discrimination task, neither of these 3-D percepts was seen, although observers did perceive that the dots were not moving with a single common speed. In fact it would be difficult to explain the present results if observers were basing their judgments on 3-D structure because it is unclear to what stimulus cue observers could have responded. If relative velocity was used to extract 3-D structure, then in those stimuli where the mean speed changed while the range of speeds was held constant, the magnitude of perceived depth and rotation speed would also be

constant. However, we know that observers can discriminate stimuli in which only the mean speed changes. If observers perceived a 3-D structure with our stimuli, one stimulus characteristic that would reliably change with the mean speed is the density of dots at particular depth planes. But if this cue was the basis of discrimination, one must ask why this same cue could not be used when only the modal speed changed since that would also change dot density at particular depth planes.

Our point is that the human visual system is remarkably versatile in its abilities to interpret motion information. The visual system seems to be able to process information in different ways depending on what task it must perform. If required to judge SFM or depth from motion, the visual system can segregate relative speeds and interpret them as representing different depths (motion parallax). But within a different context, the visual system can also integrate many different speeds together to arrive at a global mean speed. These two pieces of information, depth from motion (segregation) and global speed (integration), are not mutually exclusive in the strict sense (see footnote 6), but attending to one seems to greatly obscure the other.

In conclusion, we have shown that when a stimulus contains a distribution of speeds, our visual system is able to integrate those many different speeds to arrive at a global speed corresponding to the mean of the distribution. We can discriminate these global speeds, composed of many spatially-intermingled speed vectors, as precisely as when all dots move at the same speed implicating a considerable level of internal noise associated with the encoding of speed. Moreover, when attending to the global speed, other stimulus characteristics such as the modal speed or trajectories of individual dots, are obscured. Finally, we have shown that the details of global speed discrimination are very similar to those of global direction discrimination. The similarities between the global speed and global direction results suggest that the characteristics of the mechanisms underlying these two global processes may be equally similar.

References

Bowne, S. F., McKee, S. P. & Glaser, D. A. (1989). Motion interference in speed discrimination. *Journal of the Optical Society A., 6*, 1112-1121.

De Bruyn, B. & Orban, G. A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research, 28*, 1323-1335.

Gottsdanker, R. M. (1956). The ability of human operators to detect acceleration of target motion. *Psychological Bulletin, 53*, 477-487.

McKee, S. P. (1980). A local mechanism for differential velocity detection. *Vision Research, 21*, 491-500.

McKee, S. P. & Nakayama, K. (1988). Velocity integration along the trajectory. *Investigative Ophthalmology and Visual Science (ARVO supplement), 29*, 266.

Morgan, M. J. (1976). Pulfrich effect and the filling in of apparent motion. *Perception, 5*, 187-195.

Morgan, M. J. (1980). Analogue models of motion perception. *Philosophical Transactions of the Royal Society of London B, 290*, 117-135.

Schmerler, J. (1976). The visual perception of accelerated motion. *Perception, 5*, 167-185.

Siegel, R. M. & Andersen, R. A. (1988). Perception of three-dimensional structure from motion in monkey and man. *Nature, 331*, 259-261.

Snowden, R. J. & Braddick, O. J. (1991). The temporal integration and resolution of velocity signals, *Vision Research, 31*, 907-914.

Treue, S., Husain, M. & Andersen, R. A. (1991). Human perception of structure from motion. *Vision Research, 31*, 59-75.

Watamaniuk, S. N. J., Sekuler, R. & Williams, D. W. (1989). Direction perception in complex dynamic displays: The integration of direction information. *Vision Research, 29*, 47-59.

Watson, A. B. & Pelli, D. G. (1983). Quest: A bayesian adaptive psychometric method. *Perception & Psychophysics, 334*, 113-120.

Williams, D. W. & Phillips, G. (1986). Structure from motion in a stochastic display. *Journal of the Optical Society of America, A3*, P12.

Williams, D. W. & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24*, 55-62.

Williams, D., Tweten, S. & Sekuler, R. (1991). Using metamers to explore motion perception. *Vision Research, 31*, 275-286.

Figure Captions

Figure 1. Schematic representations of a standard and comparison stimulus in which A) the mean speed is varied while the mode is kept constant, B) the modal speed is varied while the mean is kept constant, and C) the mean is changed in one direction (increased) while the mode is shifted in the other direction (decreased). Arrows and numbers indicate the mean speeds of the distributions. These stimuli have a density of 2.26 dots/deg^2.

Figure 2. Weber fractions, $\Delta V/V$ where V is the mean speed of the distribution, averaged over observers and plotted for each stimulus group (data have been averaged over the stimulus conditions within each group). Standard errors are based on the differences between the thresholds averaged together. Note that the thresholds are similar across stimulus groups even though physically there is more consistent information regarding the change in speed as one goes from the leftmost to the rightmost stimulus group.

Figure 3. Percent correct plotted as function of percent change in modal speed. The graph plots data, averaged over observers, for four stimulus conditions: two stimulus conditions had the rectangular distribution at the slow end and the triangular distribution, of height 7 dots for Condition 1 (○) and height 14 dots for Condition 2 (□), at the fast end of the background distribution, while the other two stimulus conditions had the positions of the rectangular and triangular distributions reversed, triangular distribution of height 7 dots for Condition 3 (●) and height 14 dots for Condition 4(■). Standard errors are based on the differences between the three thresholds averaged together and each data point is the result of 600 trials (200 per observer). The data show that when only the mode of the underlying speed distribution is changed, performance stays at chance levels (50%).

Figure 4. Weber fractions, $\Delta V/V$ where V is the mean speed of the

distributions, for Gaussian speed distributions averaged over two observers and plotted as a function of the speed distributions' standard deviations in deg/sec. Standard errors are based on the differences between the thresholds averaged together. These stimuli were produced by choosing a random sample of speeds from the distribution each frame with replacement - thus stimulus variability was equal to its distribution's variance. Notice that speed discrimination does not change significantly as the standard deviation of the speed distribution increases from 0 to 1.7 deg/sec.

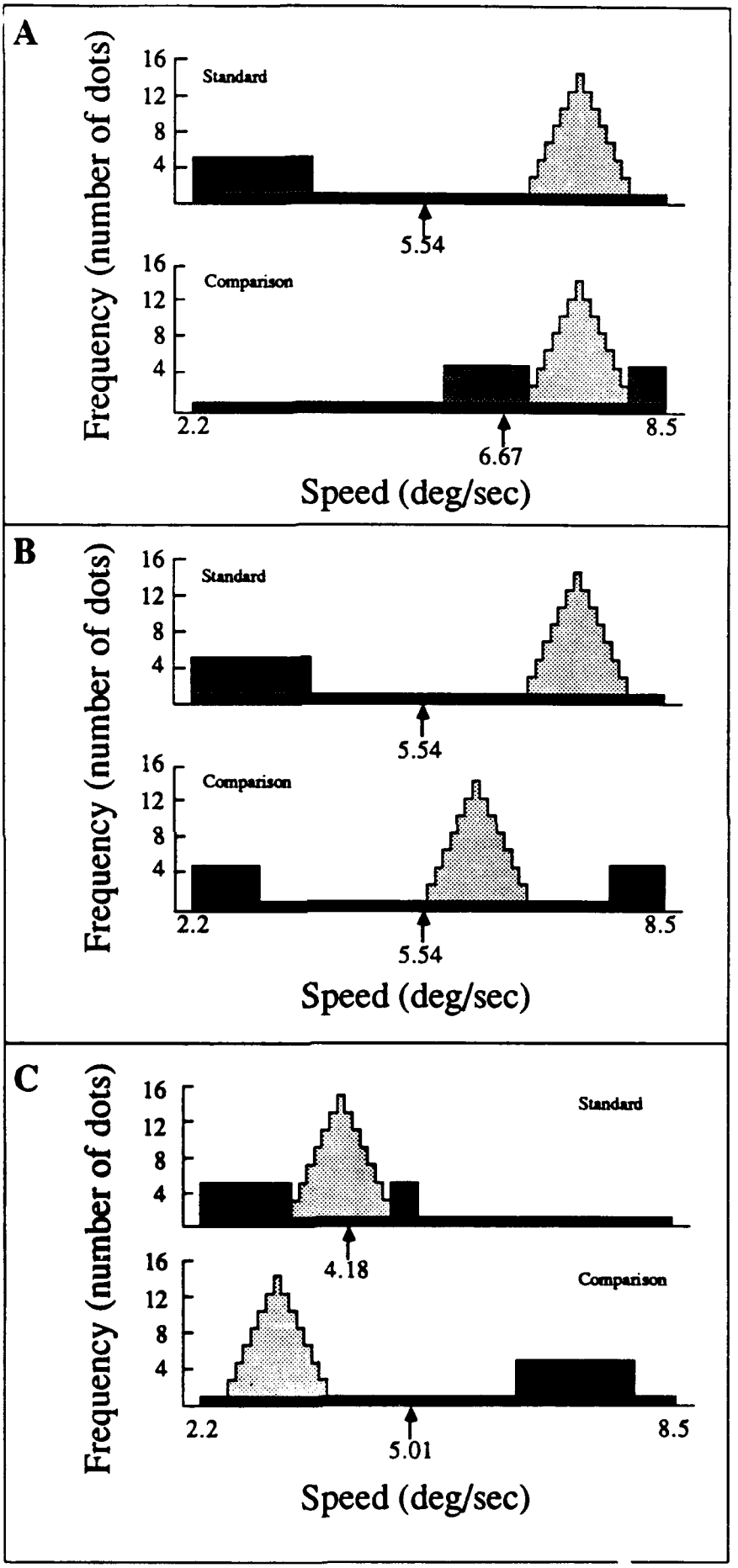
Figure 5. Schematic representation of a stimulus composed of dynamic random dots. Diagram A shows three frames of a stimulus with seven dots in which each dot chooses a new speed from the underlying distribution each frame (1-D random-walk algorithm). Diagram B shows three frames of a stimulus with seven dots in which each dot chooses a speed from the underlying distribution in the first frame and then continues at that speed for the duration of the stimulus presentation (fixed-trajectory algorithm). The figure demonstrates that the movement algorithm determines where on the screen speed vectors will be presented each frame but does not change which speeds are presented.

Figure 6. Percent correct for speed discrimination plotted as a function of the percent change in the mean speed between standard and comparison stimuli for two stimulus configurations (Mean & Mode Change Together in panel A; Mean Change & Mode Constant in panel B - see text for description of stimuli) and two movement algorithms. Data averaged over three observers (O) are plotted with those of the computer simulation (●) which used the average speed of a single randomly-chosen dot in each interval to perform the discrimination. Plus and minus one standard error are plotted on each point. Both panels show that when the 1-D random-walk algorithm is used, graphs on the left, humans perform better than the computer. When the fixed-trajectory algorithm is used, humans maintain their good performance while the computer's performance decreases precipitously. The human visual system must average speed information across many dots to maintain good performance under the fixed-

trajectory movement algorithm.

Figure 7. Percent correct for speed discrimination plotted as a function of the percent change in the mean speed between standard and comparison stimuli for two stimulus configurations (Mean & Mode Change Together top panel; Mean Change & Mode Constant bottom panel - see text for description of stimuli). Data, averaged over three observers, are plotted for three movement-algorithm conditions. Plus and minus one standard error are plotted on each point. Data from Figure 6, random-walk (O) and fixed-trajectory (●) algorithm conditions, have been replotted along with data from the 'mixed' condition (□) in which one stimulus within a trial was produced using the 1-D random-walk algorithm and the other using the fixed-trajectory algorithm. For both stimulus configurations, the three curves overlap each other considerably showing that observers were using the same viewing strategy, averaging speeds over many dots, regardless of the type of movement algorithm used.

Figure 1


A
Frequency (number of dots)
16
12
8
4
Standard
5.54
Comparison
2.2
8.5
6.67
Speed (deg/sec)
B
Frequency (number of dots)
Standard
5.54
Comparison
2.2
8.5
5.54
Speed (deg/sec)
C
Frequency (number of dots)
Standard
4.18
Comparison
2.2
8.5
5.01
Speed (deg/sec)

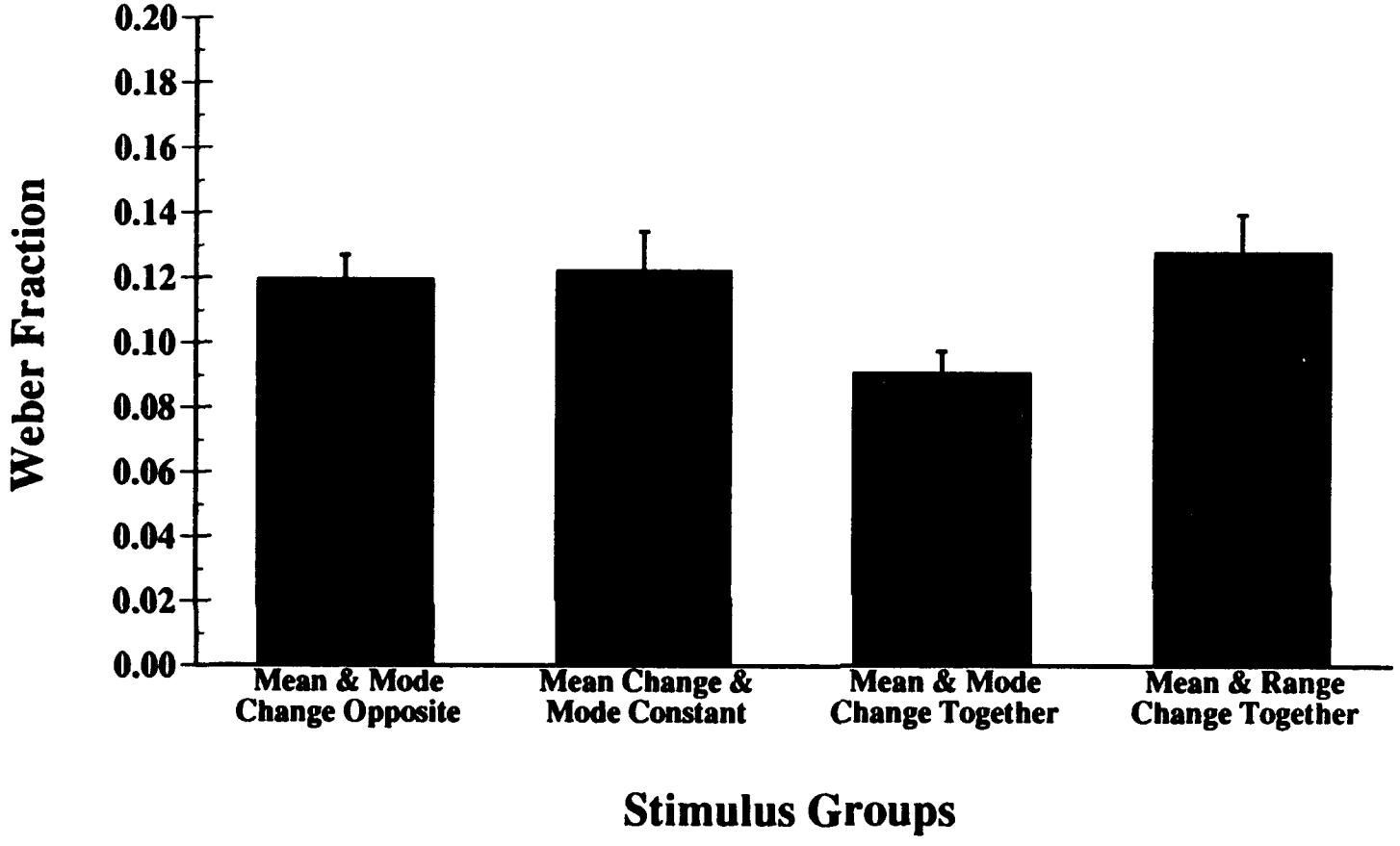

Weber Fraction
0.20
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Mean & Mode Change Opposite
Mean Change & Mode Constant
Mean & Mode Change Together
Mean & Range Change Together
Stimulus Groups

Figure 2

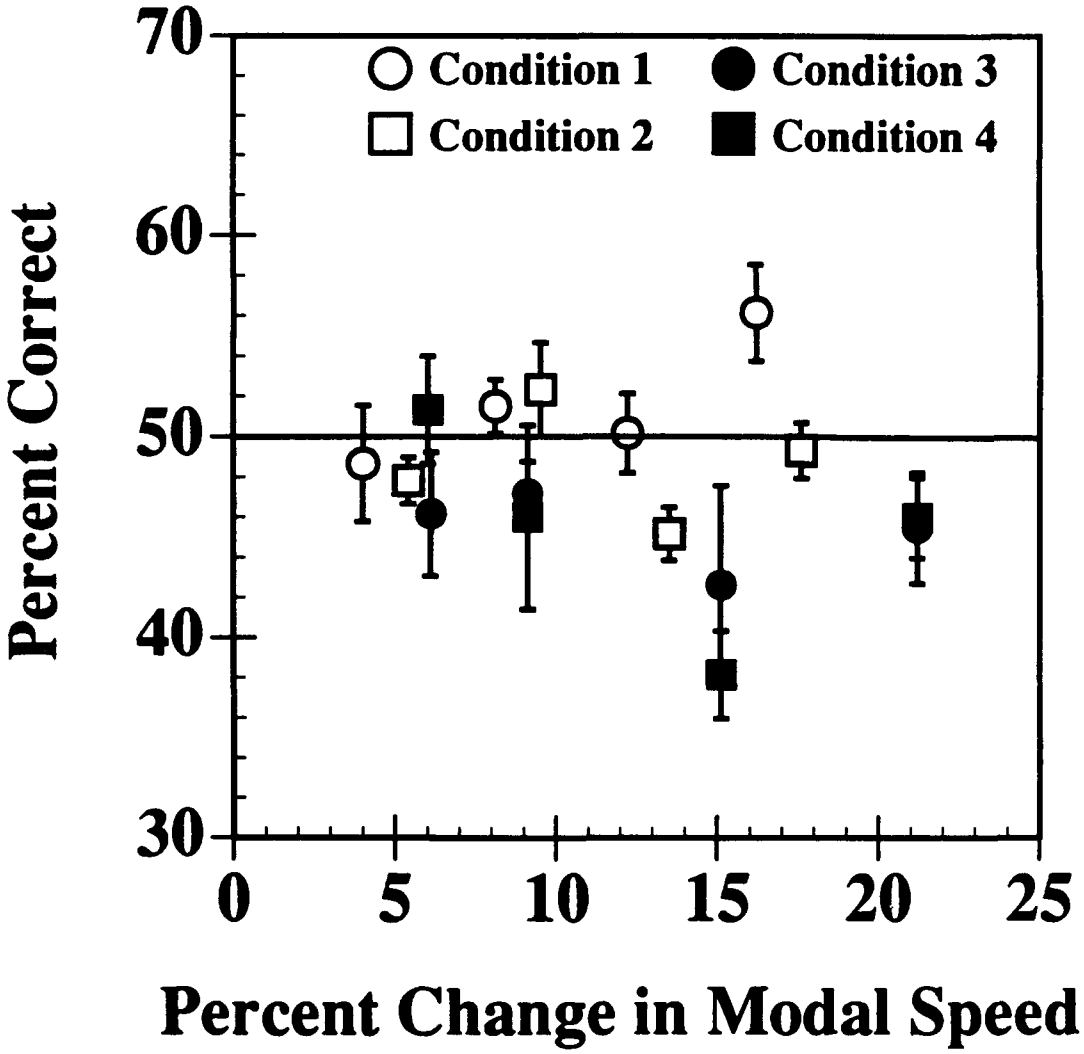
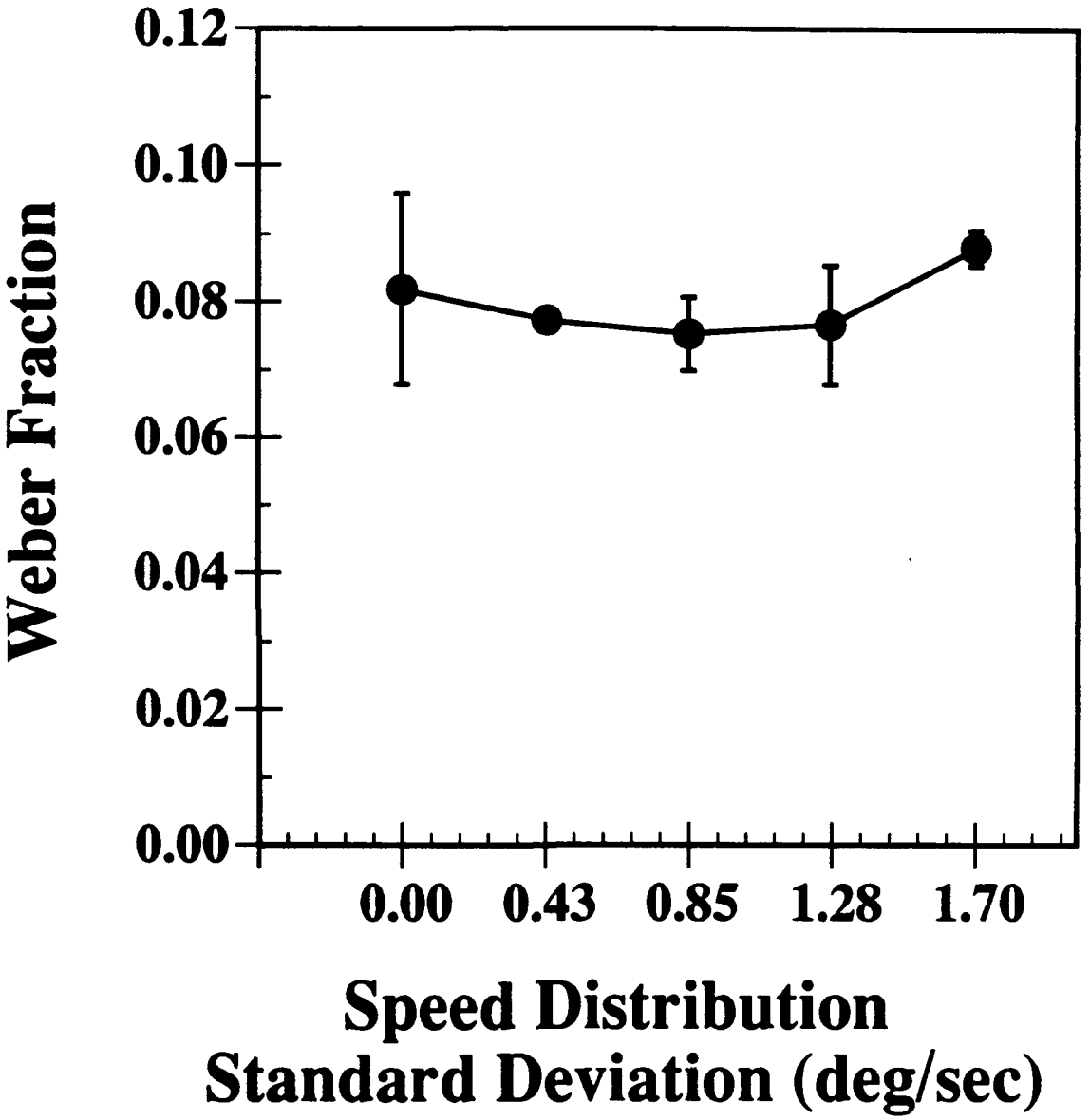
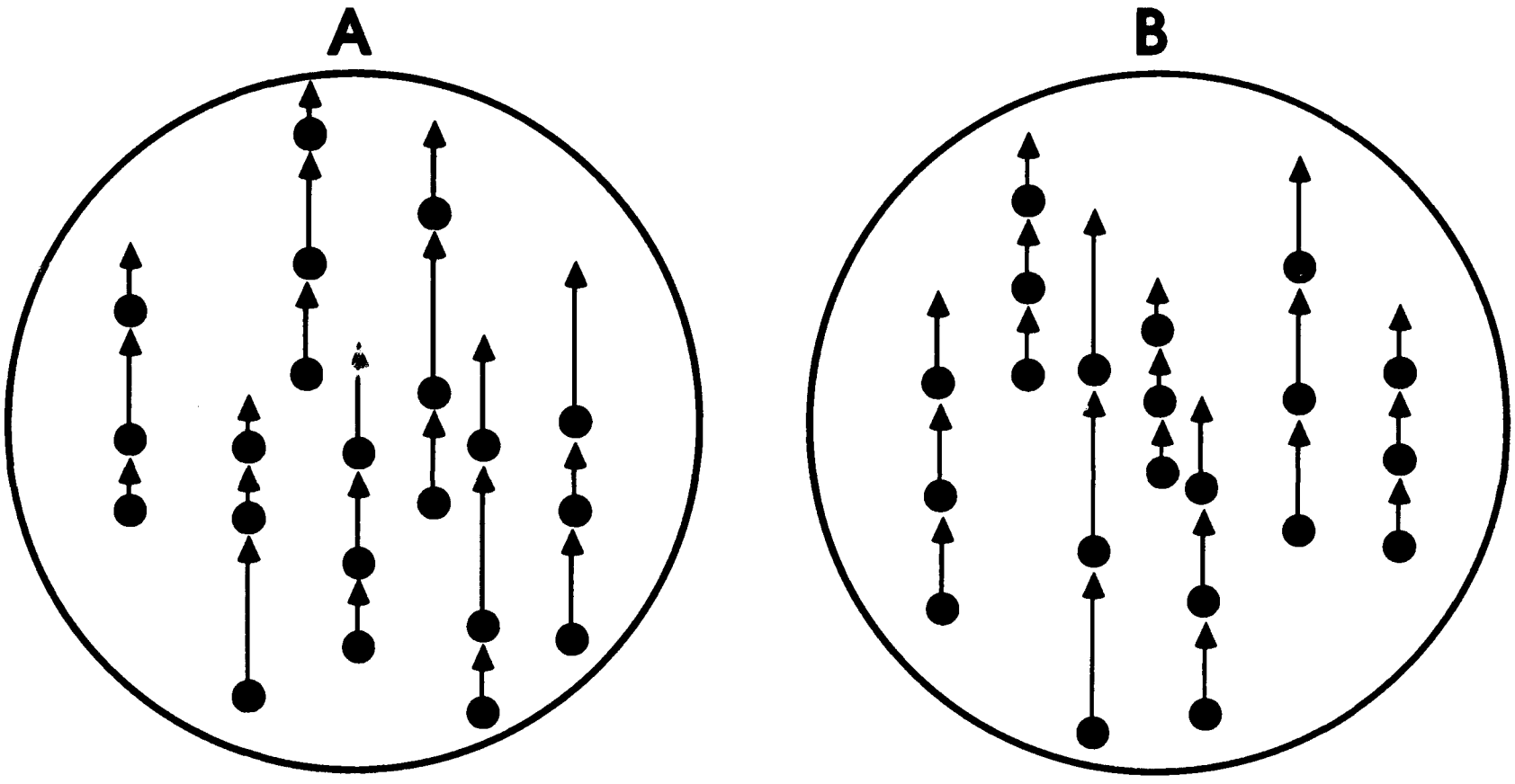

70
60
50
40
30
Percent Correct
Condition 1
Condition 3
Condition 2
Condition 4
0
5
10
15
20
25
Percent Change in Modal Speed

Figure 3

Figure 4


0.12
0.10
0.08
0.06
0.04
0.02
0.00
Weber Fraction
0.00
0.43
0.85
1.28
1.70
Speed Distribution
Standard Deviation (deg/sec)



1-D Random-Walk

Fixed-Trajectory

Figure 5

Figure 6

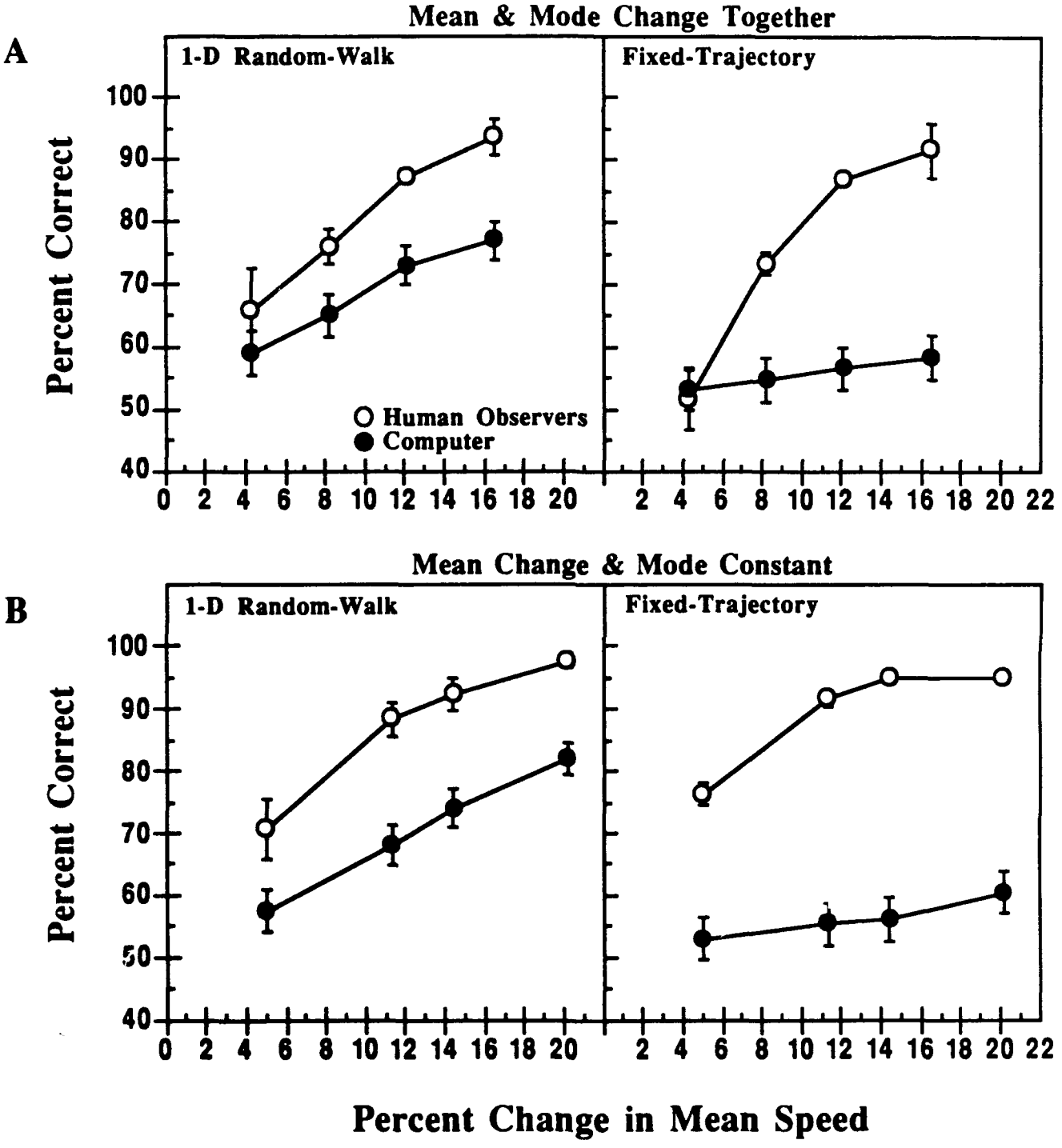
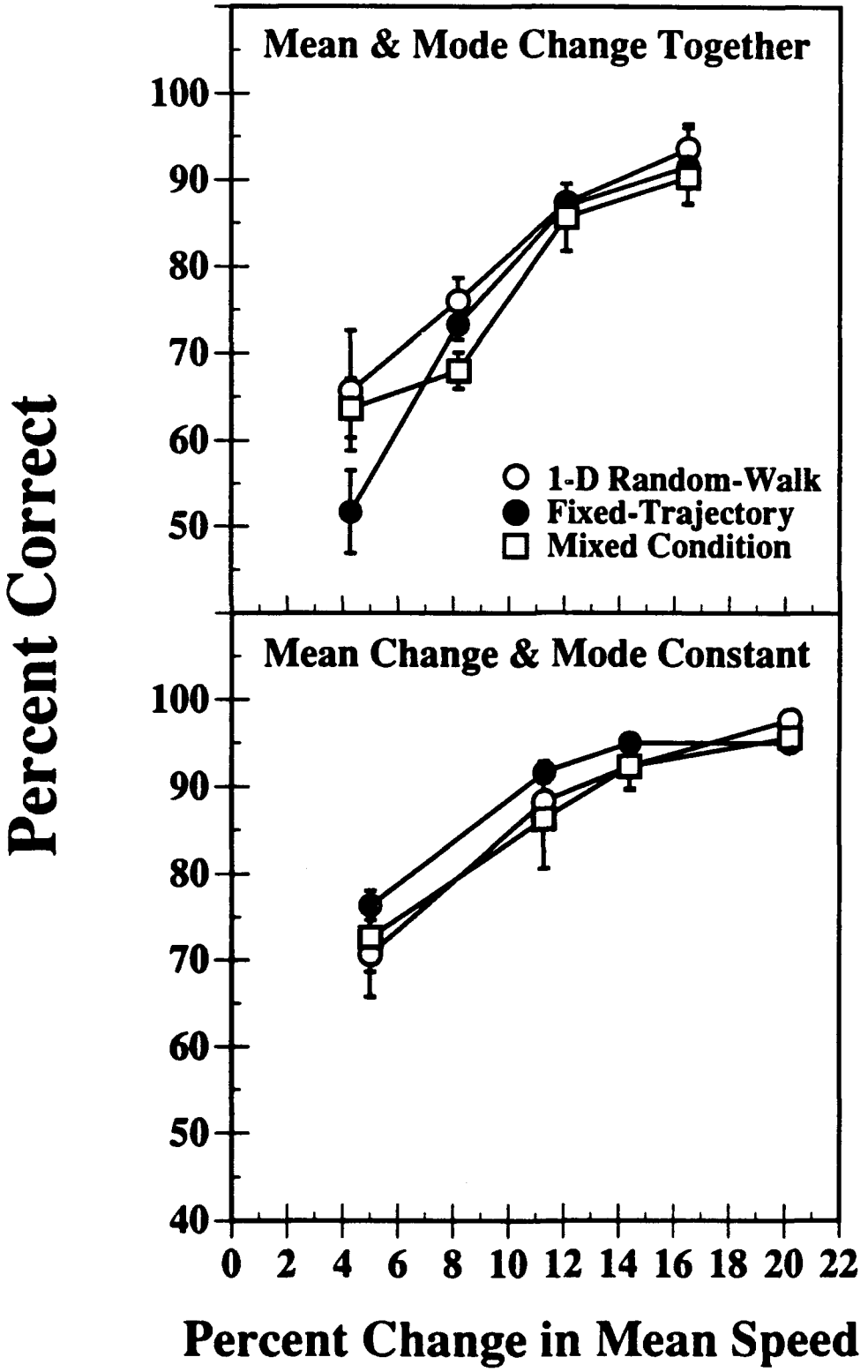

A
Mean & Mode Change Together
1-D Random-Walk
Fixed-Trajectory
Percent Correct
100
90
80
70
60
50
40
0 2 4 6 8 10 12 14 16 18 20
2 4 6 8 10 12 14 16 18 20 22
Human Observers
Computer
B
Mean Change & Mode Constant
1-D Random-Walk
Fixed-Trajectory
Percent Correct
Percent Change in Mean Speed

Figure 7


Mean & Mode Change Together
100
90
80
70
60
50
1-D Random-Walk
Fixed-Trajectory
Mixed Condition
Mean Change & Mode Constant
100
90
80
70
60
50
40
0 2 4 6 8 10 12 14 16 18 20 22
Percent Correct
Percent Change in Mean Speed

Reprinted from Nature, Vol. 337, No. 6209, pp. 734-736, 23rd February, 1989

The perception of moving plaids reveals two motion-processing stages

Leslie Welch

Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, California 94115, USA and
University of California at Berkeley, School of Optometry, Minor Hall, Berkeley, California 95720, USA

When viewed through a small aperture, the perceived motion exhibited by a long moving line or grating is ambiguous. This situation prevails because even a perfect machine could only detect motion perpendicular to a moving contour, so motion parallel to a contour is undetectable. The human visual system views the world through an aperture array—the neural receptive fields. Therefore a moving object is viewed through many small apertures and the motion within many of those apertures is ambiguous. This ambiguity may be resolved by monitoring the motion of a distinctive feature, such as a line-end or corner, and attributing to the larger object the motion of the feature. Alternatively, Adelson and Movshon[1] have suggested that moving images are processed in two stages, that is, they are first decomposed into one-dimensional components which are later recombined to generate perceived object motion. For a moving plaid, defined as the sum of two drifting gratings (Fig. 1), these alternative models generate different predictions concerning the resolution of the plaid's motion ambiguity. A feature monitor would respond to the motion of the intersections between gratings, whereas the two-stage motion processor would first decompose the plaid into its constituent gratings and subsequently recombine them to generate the perception of a moving plaid. Using speed discrimination to distinguish between the two models, I find that discrimination thresholds reflect the speed of a plaid's component gratings, rather than the speed of the plaid itself. This result supports the two-stage model. Although speed discrimination is limited by component processing, observers cannot directly access component speed. The only perceptually accessible velocity signal is generated by the second-stage pattern processing.

A plaid's speed is equal to the speed of a component grating, in the direction perpendicular to its orientation, divided by the cosine of the angle between the plaid's direction of motion and the grating component's direction of motion. Therefore the plaid always moves faster than the components: the larger the difference in component directions, the greater the difference between the component grating speed and the plaid speed. In this study, the two gratings forming the plaid have different orientations but identical spatial and temporal frequencies. As human speed discrimination varies with the speed of the target, it is appropriate to enquire whether it is the speed of a plaid's component gratings or the speed of a plaid itself that limits speed discrimination. The feature-monitor model predicts that speed discrimination is limited by the speed of the plaid, whereas the two-stage model predicts that the speed of the component gratings sets the limit.

Plaids formed by gratings at ±78° (see Fig. 1) move five times faster than their underlying component gratings. On each trial the stimulus moved at a uniform speed chosen randomly from a narrow range of speeds, and the observers were asked to judge whether it moved faster or slower than the mean speed of the narrow range. Observers (L.W. and D.T.) were provided with practice trials and feedback, and they readily learned the mean speed of the set. The smallest reliably detected change in speed was determined, and that incremental change was divided by the mean speed to calculate the Weber fraction for speed (for more complete methods see ref. 2). Expressed as a percentage, this Weber fraction is the threshold measurement reported here.

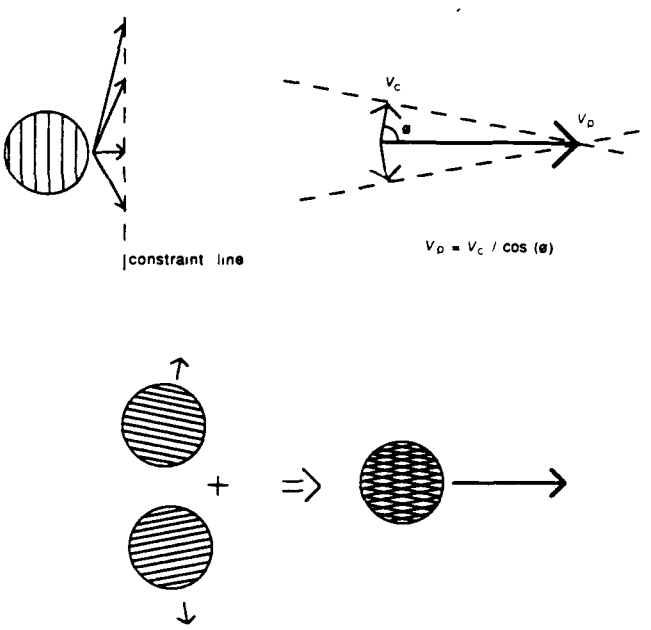


Fig. 1 The motion of a grating viewed through a stationary aperture is ambiguous: motion straight to the right cannot be distinguished from faster motion up and to the right. A family of possible vectors describing the grating motion are constrained to end on the dotted line as illustrated. Two gratings moving in different directions define two constraint lines which intersect at a position designating the velocity that is consistent with the motion of both gratings. The pattern speed is determined by the intersection of constraint lines. Illustrated here are two gratings at ±78° forming a plaid. The vector to the intersection of constraint lines indicates that the plaid's speed is much faster than the gratings' speed.

thresholds, then identical speed discrimination would be found for a grating moving at $1° s^{-1}$ and for a plaid moving at $1° s^{-1}$. Clearly the predicted relationship does not prevail. Consequently, pattern speed does not predict speed discrimination for plaids. The same discrimination data are replotted as a function of component speed in Fig. 2*b*. If the component speed predicts the thresholds, then discrimination for a grating moving at $1° s^{-1}$, would be identical to the value for a plaid moving at $5° s^{-1}$, that is, a plaid formed by gratings moving at $1° s^{-1}$. This prediction is consistent with my observations. Component speed quantitatively predicts speed discrimination for plaids.

It could be argued that the temporal frequency, namely the rate at which features pass a fixed position in space, is the same for a plaid as for its components, and that the limitation shown in Fig. 2*b* represents a generalized temporal frequency limitation on speed discrimination. Indeed, both temporal frequency and spatial frequency, the spacing between bars of a grating, affect discrimination. To study the effects of temporal frequency, I generated a plaid pattern made of high-spatial-frequency components: 12 cycles per degree for observer L.W. and 15 cycles per degree for observer D.T., but moderate temporal frequency, 6 Hz. Discrimination for the high-spatial-frequency plaid was compared with a grating characterized by the same temporal frequency and moving with the same speed as the plaid. It should be noted that the spacing of the dark bars of the control grating is identical to the spacing of the plaid intersections along any given horizontal section. Figure 3 displays the following results: discrimination for the plaid was poor, but discrimination for the grating with the identical temporal frequency was excellent. Clearly, the pattern's temporal frequency does not limit speed discrimination.

The salient features of a plaid are the intersections of the component gratings which form lighter and darker 'nodes'. If

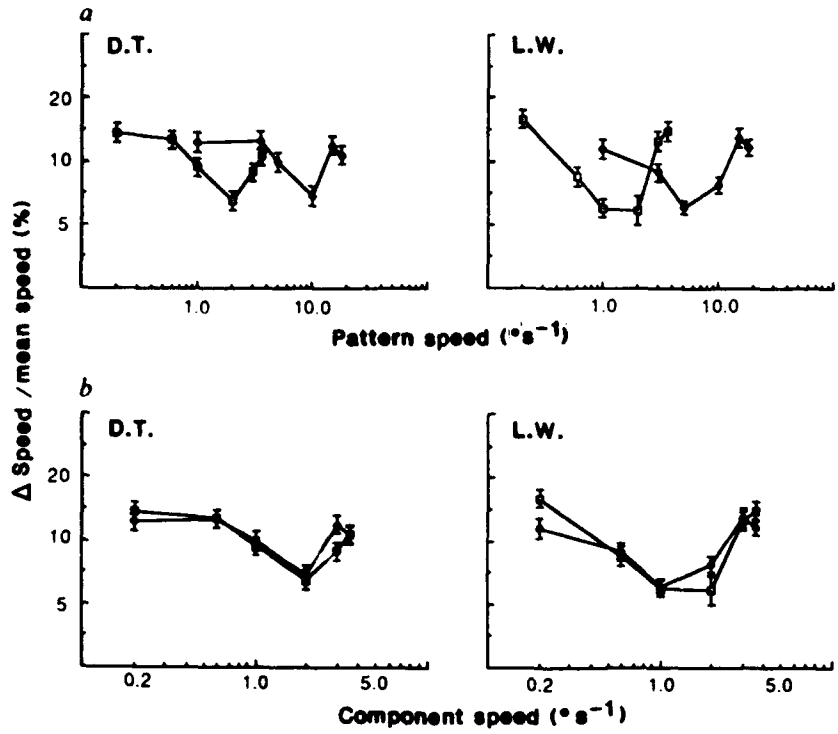


Fig. 2 *a*, Weber fraction percentages for speed discrimination are plotted as a function of pattern speed for two observers, with error bars ±1 standard error of the mean. The grating contrast was 12.5% and the average luminance 20 cd m^{-2}. The stimuli filled a 4.5° circular aperature and were presented for a variable duration between 0.4 and 0.6 s. Gratings for observers were 3.5 cycles per degree for observer L.W. and 3.0 cycles per degree for observer D.T. Open symbols indicate thresholds for single gratings and closed symbols for plaids. The grating and plaid curves do not superimpose. *b*, Same data as in *a*, replotted as a function of component grating speed. Open symbols signify single gratings and closed symbols signify plaids. The two curves superimpose.

to that for the plaid. To test this idea, I used a grid of dots with spatial positions and speed identical to the dark nodes (illustrated at the bottom of Fig. 3). Speed discrimination for the dot-grid was superior to the thresholds for the corresponding plaid (Fig. 3).

Both the present data and plaid-masking data[3] support the Adelson and Movshon two-stage model[1,4]. These results also demonstrate that the component stage is the site of limiting noise for precise speed discrimination. This does not imply that the signal at the component stage is accessible to the observer for speed discrimination. Rather, the information must pass through the second or pattern-processing stage before it becomes available to the observer. The following experiment demonstrates this constraint: discrimination for single gratings, with orientations randomized between trials, is compared with discrimination for plaids, whose components were similarly varied. Speed discrimination for a single grating was not impaired by randomized orientation, as shown in Fig. 4. Random orientation of the single gratings does not affect component speed, but varying the angle between the components of a plaid pattern introduces large variations in the pattern speed. By this procedure, the speeds of the components and the plaid are decou-

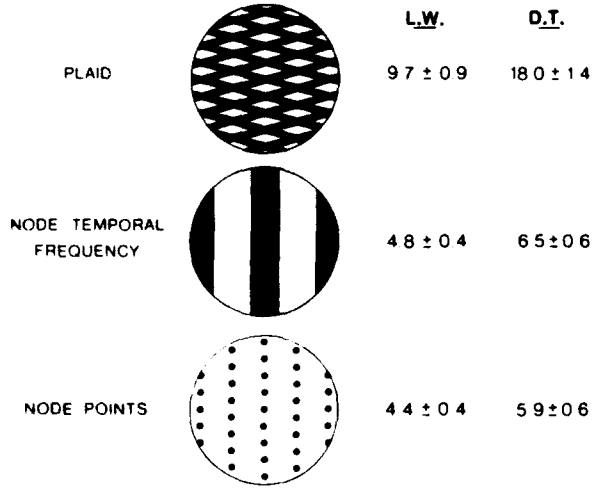


Fig. 3 Speed-discrimination thresholds, expressed as percentage of mean speed, are compared for three patterns moving with the same speed. For observer L.W., all patterns moved at 2.5° s^{-1}. The plaid was the sum of 12 cycles per degree gratings drifting at 6 Hz, and the node temporal frequency grating was 2.4 cycles per degree drifting at 6 Hz. For observer D.T., the patterns moved at 2.0° s^{-1}; the plaid was the sum of 15 cycles per degree gratings drifting at 6 Hz, and the node temporal frequency grating was 3.0 cycles per degree drifting at 6 Hz. Note that the spatial frequency of the 'node

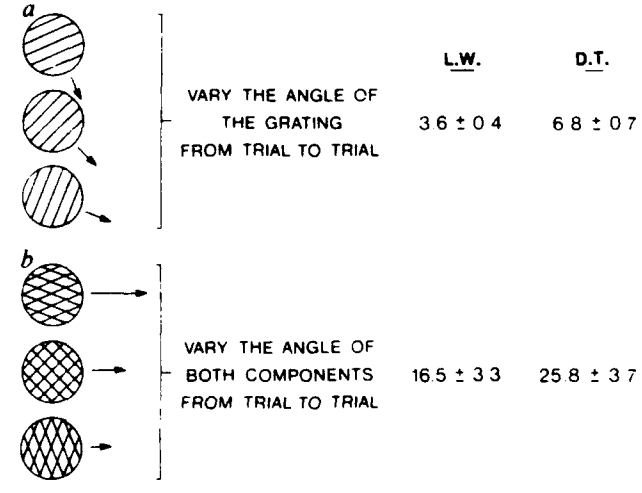


Fig. 4 *a*, A single grating at various angles; angles were chosen randomly on each trial from a set of 7 fixed angles ranging from 62° to 28°, or 45° mean. Gratings were 2 cycles per degree spatial frequency, and 5 Hz temporal frequency. Speed discrimination for this condition was good. The lower spatial frequency of these gratings results in slightly better speed discrimination thresholds than the experiment shown in Fig. 2. *b*, A plaid with both components changing angles; angles were again between 62° and 28°.

pled. Hence, on any single trial, when the components moved faster than the mean speed of the test set, the plaid may actually have moved slower than the mean, because the angle between the components was decreased on that particular trial. During these trials, observers were asked to discriminate the speed of the plaid's component gratings. The data in Fig. 4 indicate that the observers viewing the plaid were unable to judge the component's speed accurately, although the same information was available here as in the single grating instance. In conclusion, it follows that the pattern speed is the only information to which the observers could respond.

The Adelson and Movshon two-stage model receives further support from our data; component processing occurs before pattern processing, with no speed information from the component stage by-passing the pattern stage. First, a complex pattern is decomposed into its component parts and later these are combined to form a coherent moving pattern. The limiting noise is localized in the component-processing stage, which is followed by the pattern-processing stage where a velocity vector is extracted. In sum, the observer perceives the velocity of a complex pattern with a precision limited by the component-processing stage. Nonetheless, the velocity signal is available to perception only after the pattern-processing stage.

I thank Dr Ken Nakayama for suggesting the original experiment and for discussion and Drs Suzanne McKee and Samuel Bowne for help and discussion. This work was supported by the US Air Force Office of Scientific Research and the National Eye Institute.

Received 1 August 1988; accepted 9 January 1989.

1. Adelson, E. H. & Movshon, J. A. *Nature* **300**, 523-525 (1982).
2. McKee, S. P., Silverman, G. H. & Nakayama, K. *Vision Res.* **26**, 609-619 (1986).
3. Ferrera, V. P. & Wilson, H. R. *Vision Res.* **27**, 1783-1796 (1987).
4. Movshon, J. A., Adelson, E. H., Gizzi, M. S. & Newsome, W. T. *Pattern Recognition Mechanisms* (Springer, New York, 1985).

Perception, 1990, volume 19, pages 425-435

Coherence determines speed discrimination

Leslie Welch¶, Samuel F Bowne
Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, USA; ¶also at University of California, Berkeley, School of Optometry, Minor Hall, Berkeley, CA 94720, USA
Received 30 November 1989

Abstract. The visual system must determine which elements in a scene to regard as parts of a single object and which to regard as different objects. We can create stimuli that are ambiguous, ie consistent with more than one interpretation, and ask in what situations the stimulus elements are interpreted as part of a single object and when they are interpreted as multiple objects. The ambiguous stimuli in this study were moving plaid patterns—the sum of two drifting gratings with different orientations. Observers may see a rigid coherent plaid object moving in one direction, or may see two gratings moving in different directions sliding over one another. When the gratings have similar contrasts they appear to cohere and only the plaid speed is perceptually available; when the gratings have different contrasts they appear to slide and only the speeds of the gratings are perceived. Coherence thus determines what speed information is passed to higher stages of motion processing. A two-stage model of plaid motion perception is presented which agrees with the model proposed by Adelson and Movshon and extends it, detailing the relationship between coherence and speed discrimination.

1 Introduction

When an observer is presented with an ambiguous motion stimulus, such as a plaid formed by superimposing two drifting gratings with different orientations, two percepts are possible. The observer may see a rigid 'coherent' cross-hatched moving object, or see two gratings moving in different directions, transparently interpenetrating one another. Other examples of ambiguous motion stimuli that allow either rigid or nonrigid solutions are randomly moving dots on a sine-wave background (Ramachandran and Inada 1985; Ramachandran and Cavanagh 1987), displays in which dots move with velocity vectors chosen from a range of directions (Ball et al 1983; Newsome and Parré 1988), rotating line drawings (Hildreth 1984), translating line drawings (Nakayama and Silverman 1988), and hopping regions of random dots (eg Braddick 1974). All of these stimuli can be seen as moving rigidly and coherently, or as transparent, nonrigid objects with more than one motion direction. There is no physical basis for choosing one interpretation over the other, because both describe the stimulus equally well. However, Adelson and Movshon (1982) and Movshon et al (1985) have shown that when the gratings forming a plaid are similar, observers report a strong sensation of rigidity or coherence, while the transparent percept is most often seen when the gratings are dissimilar in contrast or spatial frequency (Albright and Stoner 1989). Others have shown that the gratings must be similar in color (Kooi et al 1989; Krauskopf et al 1989) and speed (Welch and Bowne 1989) as well. Coherence judgements are subjective, in that there is no correct answer, and are therefore more complicated to interpret than forced-choice judgements, such as contrast-detection thresholds. Some interpretive judgements, such as whether a cup is half-full or half-empty, seem arbitrary and idiosyncratic, and are probably unrelated to early visual processes. We wish to know whether the perceptual quality of coherence has an effect on a measure of visual performance, speed discrimination, or whether it is merely a matter of interpretation.

Several studies (Adelson and Movshon 1982; Movshon et al 1985; Welch 1989) have suggested that plaids are processed in two distinct serial stages. An illustration of a two-stage model can be found in figure 1. A plaid is decomposed into moving component gratings at the grating-processing stage and these are later used to calculate object velocity, V_p, at the plaid-processing stage. This model has been given psychophysical support by Welch (1989) who showed that speed discrimination thresholds for a coherent plaid are equal to thresholds for a component grating alone. However, this sequential model cannot be complete because two distinct percepts can be seen for superimposed gratings, coherence and sliding, as mentioned above. There must be some sort of decision process that determines which of the two possibilities is seen. Regardless of whether the gratings cohere or slide, information from both processing stages may be available for speed discrimination. A question remains: are all three velocity sectors, V_1, V_2, and V_p, passed on to higher stages of processing, or is some information lost at the plaid-processing stage? Welch (1989) showed that grating speed information is lost in a coherent plaid, while plaid speed estimates remain precise. This result suggests that coherence determines which velocities are passed on to later processing stages: V_p for coherent plaids, or the grating velocities V_1 and V_2 for incoherent plaids. The purpose of the present study is to further test the hypothesis that coherence determines what speed information is extracted from a plaid.

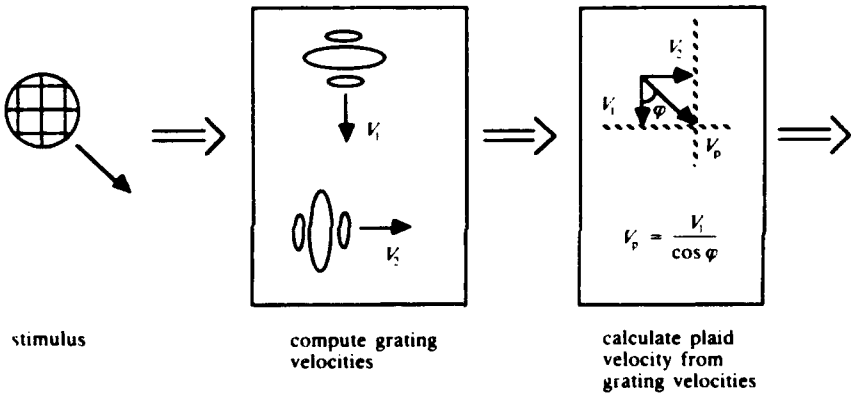


Figure 1. Adelson-Movshon model of a two-stage motion processor. The plaid is decomposed into oriented, moving gratings by the first stage. At the second stage, the grating speeds are used to calculate the plaid speed in a way that is consistent with the equation $V_p = V_1/\cos\varphi$, also termed the intersection of constraints (Fennema and Thompson 1979; Adelson and Movshon 1982).

2 General methods

An Innisfree Picasso pattern generator, under computer control, was used to display drifting sinusoidal gratings on an HP1332A oscilloscope with a P31 phosphor. Two gratings of different orientations could be superimposed by alternating between them so that each had a refresh rate of 50 Hz. The spatial frequency, temporal frequency, orientation, and contrast of the gratings could be controlled independently. An opaque circular aperture of 4.5 deg visual angle was placed on the oscilloscope face. The average screen luminance, measured with a Pritchard photometer, was 18 cd m^{-2}.

Stimuli were presented for 0.4 to 0.6 s with the duration randomized from trial to trial. The overall orientation of the stimulus pattern was randomized around 360° in approximately 6° steps. Observers viewed the screen at a 114 cm distance with natural pupils in a dimly lit room. Both authors and a well-practiced third subject

served as observers. All had good acuity for this distance with the appropriate spectacle correction if needed. A small dark fixation point was placed at the center of the screen to help reduce eye movements.

For the speed discrimination tasks we used the method of single stimuli in which one speed chosen randomly from a set of several speeds was shown on each trial [see McKee and Welch (1985) for more complete details]. The observer judged after each trial whether the speed was faster or slower than the mean speed in that block of trials. Observers were given practice trials and feedback to build up their internal representation of the mean speed. Responses were recorded by the computer and scored as the percentage of fast responses for each speed tested, which yielded a psychometric function going from near 0% to near 100%. This function was fit to a cumulative normal curve by probit analysis. Threshold is defined as the speed increment which yields a shift from 50% to 75% on the probit curve, corresponding to a d' of 0.675. Each threshold was based on at least 200 trials. Probit analysis was also used to estimate the standard error which is shown as the error bars on the graphs.

Speed was manipulated by changing the grating temporal frequency. These experiments do not distinguish between temporal frequency discrimination and speed discrimination, but McKee et al (1986) have argued that observers actually discriminate speed rather than temporal frequency.

3 Coherence experiment

3.1 *Methods*

In the first experiment, observers were shown two superimposed gratings: one had a contrast of either 5% or 15%, in separate blocks, and the other had a contrast that varied on a trial-by-trial basis from 1% to 37.5%. The gratings were always oriented 90° apart but their actual orientations varied from trial to trial. That is, they could be at 0° and 90°, or 6° and 96°, or any other combination whose difference is 90°. Observers responded after each trial whether the gratings appeared as a coherent plaid or not. Judgements were recorded by the computer and are reported here as the percentage of times the observer responded "coherent" as a function of the contrast of the variable grating. Each data point is based on at least 30 trials. Error bars were calculated from the observed probability of coherence (P) and the number of trials (N) with the aid of the standard formula for the binomial distribution,

$$\Delta P = [P(1-P)/N]^{1/2} .$$

3.2 *Results*

The percentage of trials on which the observers saw a coherent plaid is plotted as a function of the contrast of the variable grating, C_2, in the upper graphs in figures 2 and 3. The contrast of one grating, C_1, was at 5% (figure 2) or 15% (figure 3) as indicated by the arrows, while that of the second grating, C_2, was randomly chosen on each trial from a set of five values as shown on the x axis. The data describe an inverted-U-shaped function with the greatest coherence when the two gratings have the same contrast; in the upper graphs of figure 2 the functions peak at about 5% contrast and in the upper graphs of figure 3 the functions peak at about 15% contrast. Coherence decreases when the variable-contrast grating, C_2, is different from the constant-contrast grating, C_1, whether it is higher or lower in contrast, in agreement with previous results (Adelson and Movshon 1982; Movshon et al 1985). It should be noted that coherence and sliding are not all-or-none phenomena; the gratings can appear to stick together partially, but the observers are required to make a binary decision, either coherent or not. They accomplish this by adopting some criterion level of coherence and calling everything that sticks more than that coherent, and

everything that sticks less than that transparent. This means that the actual percent-coherent numbers are criterion-dependent and therefore less reproducible between observers than the general shape of the curves. The coherence graphs in figures 2 and 3 both show that observer SB judged the gratings to be less coherent in general than observers SPM and LW. The difference may reflect a real difference in the plaid appearance between the observers, or SB may simply be biased to respond "coherent" less often. This method cannot distinguish between these two possibilities.

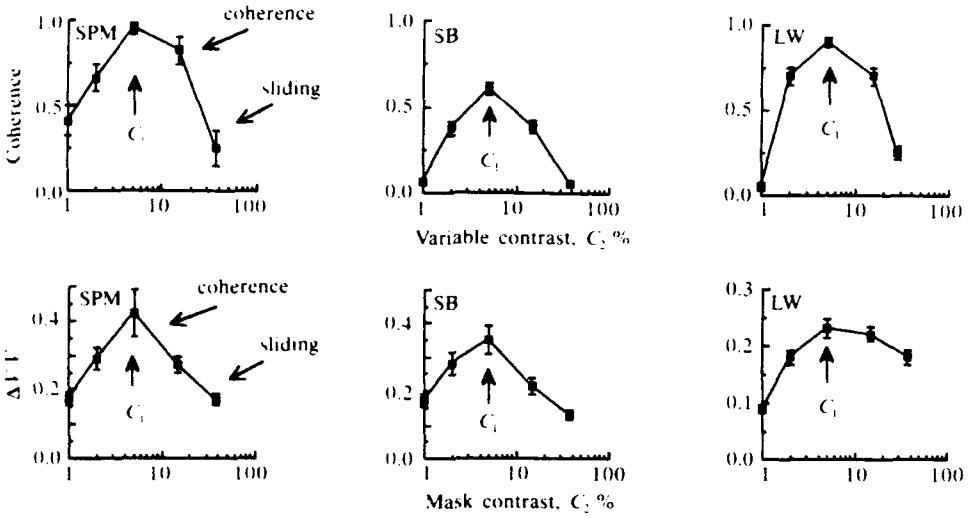


Figure 2. The upper curves show the effect of contrast on coherence. One grating, C_1, was always 5% contrast as indicated by the arrows. The other grating, C_2, varied in contrast as shown on the x axis. The vertical axis shows the percentage of displays judged to be coherent. The gratings were judged most coherent when their contrasts were equal. The lower curves show the just noticeable difference in target speed as a function of mask grating contrast, C_2. Target contrast, C_1, was 5% as indicated by the arrows. Grating speed discrimination thresholds were higher when the contrasts of the two gratings were similar (coherent), and thresholds decreased when the contrasts increasingly differed (sliding).

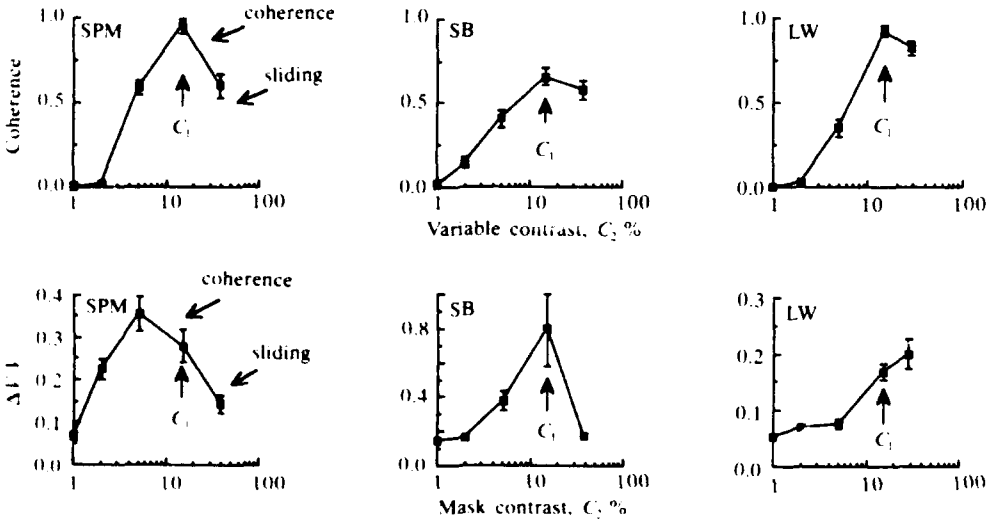


Figure 3. Coherence judgement and speed discrimination thresholds similar to those in figure 2 but for C_1 = 15%.

4 Grating speed experiment

4.1 *Methods*

The purpose of the speed discrimination experiments was to determine the relative precision of grating speed perception and plaid speed perception. We therefore developed two stimuli in which the grating and plaid speeds were 'decoupled', that is, were not proportional to each other. Plaid speed is determined by the speed of the gratings and the angle between the plaid motion direction and the grating motion direction.

The first decoupling method was to increment the speed of one of the gratings while decrementing the speed of the other, orthogonally oriented, grating (see figure 4). The speed of one grating increases by the same amount as the speed of the other grating decreases so the absolute value of the speed changes is correlated trial by trial. The sign of the speed shifts for the gratings is opposite, thus their speeds are 'anticorrelated'. This results in no significant change in the plaid speed[1] The 'anti-correlation' of the speeds of the two gratings does result in a shift in the direction of the plaid motion, but direction-of-motion information was obscured by randomizing the overall orientation of the stimulus trial by trial. 500 ms before each trial the observer was shown a low-contrast 'cue' grating which indicated the orientation of the test grating, C_1. The observer was asked to judge the speed of the cued grating and to ignore the other grating, C_2. This can be thought of as a masking paradigm where the cued grating, C_1, is the test grating and the other, C_2, is the mask.

4.2 *Results*

The observers' task was to discriminate the speed of one 'test' grating, C_1, while ignoring the other, variable-contrast 'mask' grating, C_2. The lower graphs in figures 2 and 3 show that speed discrimination thresholds are high when the contrasts of the

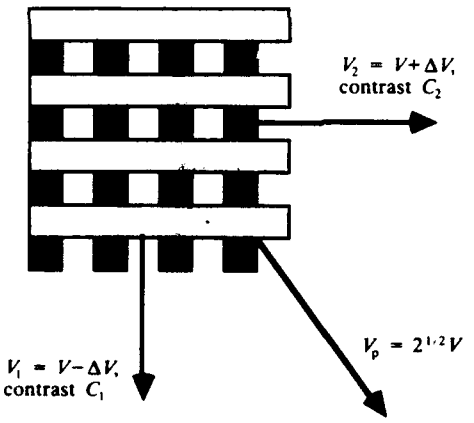


Figure 4. The grating speed discrimination stimulus is a plaid formed by two orthogonally oriented gratings. When the speed of one grating is incremented, the speed of the other is decremented by the same amount, resulting in no significant change in plaid speed[1]. The resulting shift in plaid direction of motion was obscured by randomizing the overall plaid orientation. One grating was designated the test, C_1, and its orientation was indicated by a low-contrast cue grating before each trial. Observers judged the speed of the test grating, C_1, while ignoring the other, masking grating, C_2. The plaid speed provides no information useful for this task, so the observers are forced to judge the grating speed.

[1] The exact plaid speed is $2^{1/2}V[1+(\Delta V/V)^2]^{1/2}$ which differs slightly from $2^{1/2}V$ but the residual term cannot be used to tell positive ΔV (fast) from negative ΔV (slow), and therefore cannot help the observer discriminate between them.

two gratings are similar, the same conditions that appeared more coherent in the first experiment as indicated in the upper graphs in figures 2 and 3. The speed functions do not always peak precisely at 5% or 15% contrast, but they follow the coherence functions of the same observer reasonably well. Speed discrimination thresholds were lower when observers reported less coherence, when the grating contrasts were less similar. The data from the first and second experiments parallel each other and confirm the prediction that grating speed discrimination depends on coherence.

5 Plaid speed experiment

5.1 *Methods*

Another experiment was designed to decouple grating and plaid speeds in a complementary way. If the two gratings are 90° apart in orientation as in the top panel of figure 5, the plaid speed is $2^{1/2}$ times the grating speed. Larger angles change this relationship; for example, two gratings 157° apart, as in the bottom panel of figure 5, produce a plaid moving 5 times as fast as the gratings. We selected a set of seven relative orientations from 60° to 120° and seven plaid speeds from 3.5 to 6.5 deg s^{-1} and then calculated the grating speeds required to produce each plaid speed at each angle for a total of 49 patterns. On each trial a random pattern was shown and the observers were instructed to judge the plaid speed. Because of their random orientations and random speeds, the speeds of the gratings did not provide accurate plaid speed information. Therefore, in this experiment we measure the precision with which observers can perform the plaid speed calculation described in figure 1. Since only a limited range of angles was used, the grating and plaid speeds were still partially correlated such that speed discrimination thresholds of $\Delta V/V = 14\%$ were possible when only the grating speed was used as a decision variable. However, as we shall show below, thresholds below this value were found for coherent plaids, indicating the action of pattern-speed selective mechanisms.

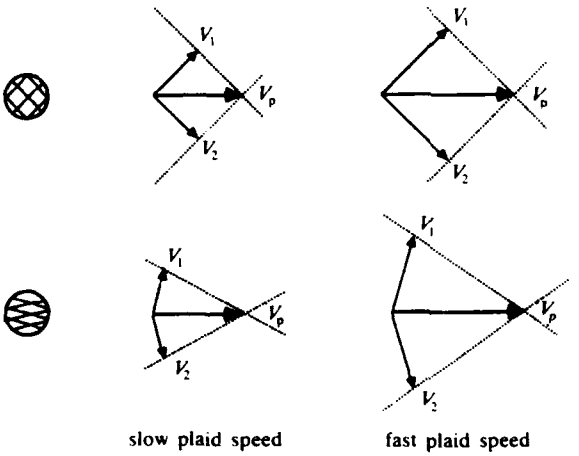


Figure 5. Plaid speed discrimination task. When two gratings 90° apart in orientation are superimposed, the plaid speed is $2^{1/2}$ times the grating speeds (top panel). If the grating velocities are more than 90° apart as in the bottom panel, slower grating motion produces the same plaid speed. By randomly mixing several orientations, we produced a stimulus in which the grating speed does not provide accurate information about the plaid speed thus forcing the observers to judge plaid speed.

5.2 *Results*

In this last experiment, grating and plaid speeds were decoupled in a manner complementary to what was done in the second experiment. Observers were asked to judge the speed of the plaid while ignoring the speeds of the gratings. In this situation with 'good' plaid-speed information and random grating speed, the pattern of results is expected to be the opposite of the results from the second experiment. This is because the 'good' plaid information cannot be perceived unless the gratings are similar in contrast and are perceived as a coherent plaid. Precise speed discrimination is expected when the contrasts of the two gratings are similar, and poor speed discrimination is expected when the contrasts are quite different. The prediction is confirmed by the data in figure 6, which are threshold speed increments plotted as a function of the contrast of grating C_2. The contrast of grating C_1 was always at 5%, as indicated by the arrows. The data form a U-shaped function with a broad minimum where the contrasts of the two gratings are similar. The horizontal line labelled 'single grating' shows the speed discrimination thresholds if one of the gratings is completely turned off. The line therefore gives an indication of the best speed discrimination possible without any knowledge of plaid speed. The points below this line demonstrate the operation of pattern-speed sensitive mechanisms. Pattern-speed mechanisms are only useful when the contrasts are similar, that is, when the gratings cohere. This second, very different method also results in speed discrimination data which correlate with the subjective impression of coherence.

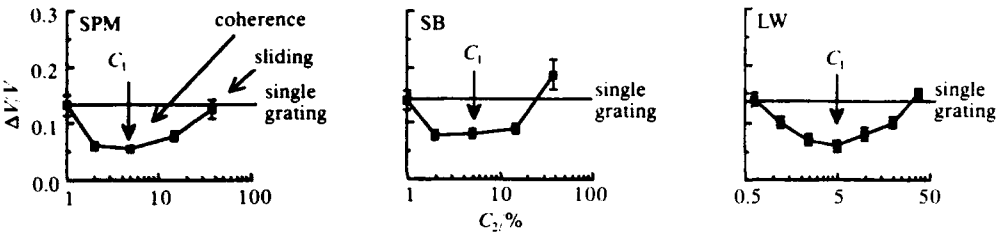


Figure 6. The effect of contrast on plaid speed discrimination. The contrast of one grating, C_1, was C_1 = 5% as indicated by the arrows while the contrast of the other, C_2, varied as shown on the x axis. Plaid speed discrimination thresholds were higher when the contrasts of the gratings were quite different (sliding), and thresholds decreased when the contrasts were similar (coherence). The horizontal line labelled 'single grating' shows the plaid speed discrimination threshold when only one grating was visible, and thus shows the best speed discrimination possible without any knowledge of plaid speed. The points below this line demonstrate the operation of pattern-speed sensitive mechanisms. Pattern-speed mechanisms are only useful when the contrasts are similar, that is, when the gratings cohere.

6 Conclusions

Given the results of the preceding experiments, the Adelson–Movshon model illustrated in figure 1 can be further developed. The serial motion-processing model has different speed information at two separate stages. This difference in information can be used to study the rules governing whether two differently oriented gratings are interpreted as a single coherent object. The first-stage mechanisms are oriented and respond to the superimposed gratings independently. Their outputs give information about the contrasts, orientations, and spatial and temporal frequencies of the gratings.

The information from the first-stage mechanisms can be used in two different ways. A minimalist model compares first-stage information directly at an early decision stage which determines whether the gratings will appear to cohere into a single object or will appear to be independent objects (figure 7). If the gratings cohere into a plaid, the plaid velocity is calculated and the grating velocities are lost. If the gratings do not

cohere, the grating velocities are calculated and the plaid velocity is lost. The output of the first-stage mechanisms is not strictly speed information in this model but rather spatial- and temporal-frequency information. A model with a later decision stage would also work and may be more likely physiologically. The first-stage information is used to calculate the velocities of the gratings and of the plaid, and the coherence decision stage determines which velocities are retained (figure 8). Either way, coherence determines whether grating velocities or plaid velocity will be available to perception and information is lost as a result of the coherence decision. Both of these models are consistent with the data presented in this paper.

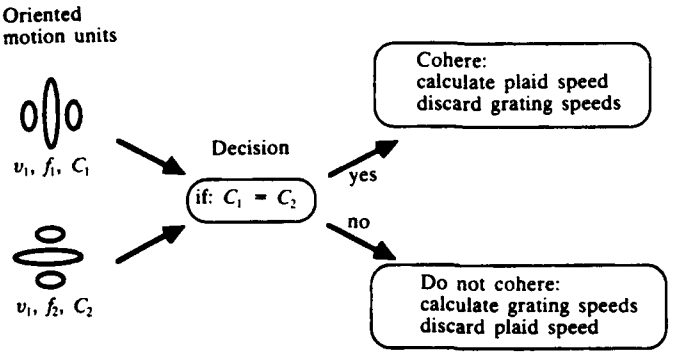


Figure 7. Illustration of an early coherence decision two-stage motion model. The two gratings excite different first-stage motion units which encode the grating orientations, contrasts (C_1 and C_2) and spatial (v_1 and v_2) and temporal (f_1 and f_2) frequencies. The contrast signals from both motion units are compared by a decision process which determines whether the gratings will be seen as a single plaid or as independent gratings. If the contrasts are similar, the decision is for a single coherent plaid and the second stage calculates the plaid velocity. If the contrasts are different, the decision is for transparency and the second stage calculates the grating velocities.

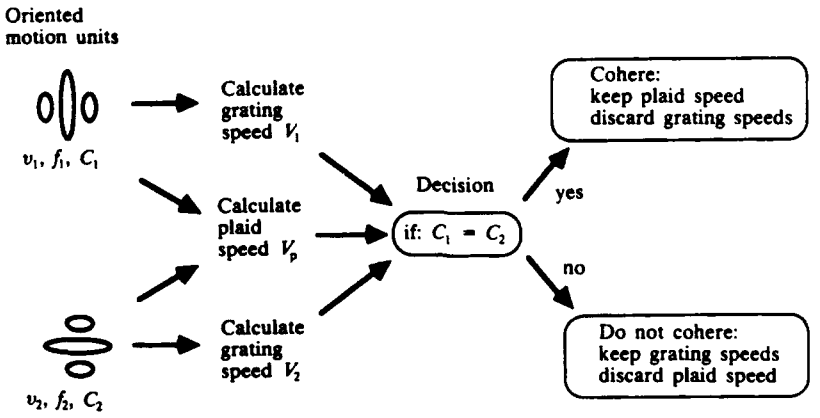


Figure 8. Illustration of a late coherence decision two-stage motion model. The first stage is the same as for the other model but the information from the motion units is used by the second stage to calculate grating and plaid velocities immediately. The contrast signals from both motion units are compared by a decision process which determines whether the gratings will be seen as a single plaid or as independent gratings. If the contrasts are similar, the decision is for a single coherent plaid, the plaid velocity is passed on and the grating velocities are lost. If the contrasts are different, the decision is for transparency, the grating velocities are passed on and the plaid velocity is lost.

7 Discussion

The speed discrimination results reveal an important point about the coherence data. Coherence is not a matter of conscious interpretation, but rather an involuntary process whereby low-level stimulus variables such as relative contrast determine what speed information is retained. Other work using coherence judgements (Adelson and Movshon 1982; Movshon et al 1985; Albright and Stoner 1989; Ho and Berkley 1989; Kooi et al 1989; Krauskopf et al 1989) should therefore be interpreted as measuring the properties of a low-level object-segregation process.

If the gratings cohere, the models in figures 7 and 8 predict a perfect plaid velocity calculation, obeying the line-of-constraints computation shown in figure 1. However, Kooi et al (1988) and Stone et al (1988) found that the direction the plaid appears to move can be influenced by the relative contrasts of the two gratings even when the gratings are seen as moving coherently. This deviation is consistent with the visual system interpreting lower-contrast gratings as moving slower than higher-contrast gratings. Thompson (1982) has found that, for some temporal frequencies, lower-contrast gratings appear to move more slowly than otherwise identical gratings of higher contrast. Another deviation from the intersection-of-constraints solution has been described by Ferrera and Wilson (1988) for asymmetric plaids made of gratings whose velocity vectors are both on the same side of the plaid velocity vector. The perceived direction of asymmetric plaid motion is biased toward the direction of the grating motions. These data do not challenge the idea of two-stage motion processing, but rather point out that the visual system does not extract object velocity perfectly. Our results indicate that, although the direction may not be perfectly calculated, the plaid speed is quite precisely determined, leading to plaid speed discrimination thresholds of 5% to 7% as shown in figure 6.

The only psychophysical challenge to the two-stage motion processing idea comes from Gorea and Lorenceau (1989). They have suggested that there are two distinct motion processes, one that responds to the plaids, such as the Perrone model (1990), and another that responds separately to the gratings, but, for the sake of parsimony, argue that the two are parallel rather than serial. We submit, given the present data and the data of Adelson and colleagues, that the more parsimonious conclusion is for a serial two-stage motion process.

The models in figures 7 and 8 are oversimplified on two bases. They consider only the cases of total coherence and total sliding, but coherence and transparency are not all-or-none phenomena. Subjectively, there is a range of conditions when the gratings cohere partially, as mentioned in the first experiment. In this situation the visual system seems to have partial speed information from both the individual gratings and the plaid, though in a degraded form as speed discrimination thresholds show intermediate values between those expected for grating speeds and plaid speeds when coherence is partial. The models also consider contrast as the only variable determining coherence. However, others have shown that additional variables such as spatial frequency (Adelson and Movshon 1982; Albright and Stoner 1989; Welch and Bowne 1989), color (Kooi et al 1989; Krauskopf et al 1989), and temporal frequency (Welch and Bowne 1989) also affect coherence. An improved model could be developed using the outputs of oriented motion sensors, but such models are beyond the scope of this paper. Although models of object motion have been proposed (Marr and Ullman 1981; Bülthoff et al 1989; Grzywacz and Yuille 1990; Perrone 1990; Sereno 1989; Wang et al 1989) they do not explicitly include a coherence decision and therefore require modification. Heeger's model of object motion (1987) has an explicit coherence decision but, as he points out (footnote 32), his model does not extract the correct plaid speed when the angle between the gratings is not 90°. This

means that his model could not predict the discrimination thresholds in figure 6 for plaids with variable grating angles.

The visual system must determine what contours belong to the same object in order to make possible appropriate interactions with the environment. Once a decision has been made about what is an object, there is no reason for the visual system to burden itself with irrelevant information like the local speeds of contours. The coherence decision is an example of an information gate that can be observed with the grating and plaid speed decoupling methods employed in this study. The important finding is not that speed discrimination is another method for measuring coherence, but that a decision rule is in place to determine whether coherence or sliding is seen and this decision process gates velocity information.

Acknowledgements. We thank Dr Suzanne McKee for her invaluable help and discussion. This work was supported by AFOSR-89-0035.

References

Adelson E H, Movshon J A, 1982 "Phenomenal coherence of moving visual patterns" *Nature (London)* **300** 523 - 525

Adelson E H, Bergen J R, 1985 "Spatio-temporal energy models for the perception of motion" *Journal of the Optical Society of America A* **2** 284 - 299

Albright T D, Stoner G R, 1989 "Motion perception survives figural cue heterogeneity" *Supplement to Investigative Ophthalmology and Visual Science* **30** 74

Anderson S J, Burr D C, 1985 "Spatial and temporal frequency selectivity of the human motion detection system" *Vision Research* **25** 1147 - 1154

Ball K, Sekuler R, Machamer J, 1983 "Detection and identification of moving targets" *Vision Research* **23** 229 - 238

Braddick O J, 1974 "A short range process in apparent motion" *Vision Research* **14** 519 - 527

Bülthoff H, Little J, Poggio T, 1989 "A parallel algorithm for real-time computation of optical flow" *Nature (London)* **337** 549 - 553

Burr D C, Ross J, Morrone M C, 1986 "Seeing objects in motion" *Proceedings of the Royal Society of London B* **227** 249 - 265

Fennema C L, Thompson W B, 1979 "Velocity determination in scenes containing several moving objects" *Computer Graphics and Image Processing* **9** 301 - 315

Ferrera V P, Wilson H R, 1988 "Perceived direction of moving 2D patterns" *Supplement to Investigative Ophthalmology and Visual Science* **29** 264

Gorea A, Lorenceau J, 1989 "Motion perception in compound stimuli is 'blob'-dependent" *Supplement to Investigative Ophthalmology and Visual Science* **30** 388

Grzywacz N M, Yuille A L, 1990 "A model for the estimate of local image velocity by cells in the visual cortex" *Proceedings of the Royal Society of London B* **239** 129 - 161

Heeger D J, 1987 "Model for the extraction of image flow" *Journal of the Optical Society of America A* **4** 1455 - 1471

Hildreth E C, 1984 *The Measurement of Visual Motion* (Cambridge, MA: MIT Press)

Ho W A, Berkley M A, 1989 "Dichoptic perception of coherent plaid movement is spatial frequency dependent" *Supplement to Investigative Ophthalmology and Visual Science* **29** 75

Kooi F L, De Valois R L, Wyman T K, 1988 "Perceived direction of moving plaids" *Supplement to Investigative Ophthalmology and Visual Science* **29** 265

Kooi F L, De Valois K K, Grosof D H, Switkes E, 1989 "Coherence properties of colored moving plaids" *Supplement to Investigative Ophthalmology and Visual Science* **30** 389

Krauskopf J, Farell B, Movshon J A, 1989 "Phenomenal coherence of moving chromatic gratings" *Supplement to Investigative Ophthalmology and Visual Science* **30** 389

Marr D, Ullman S, 1981 "Directional selectivity and its use in early visual processing" *Proceedings of the Royal Society of London B* **211** 151 - 180

McKee S P, Welch L, 1985 "Sequential recruitment in the discrimination of velocity" *Journal of the Optical Society of America A* **2** 243 - 251

McKee S P, Silverman G H, Nakayama K, 1986 "Precise velocity discrimination despite random variations in temporal frequency and contrast" *Vision Research* **26** 609 - 619

Movshon J A, Adelson E H, Gizzi M S, Newsome W T, 1985 "The analysis of moving visual patterns" in *Pattern Recognition Systems* Eds C Chagas, R Gattass, C G Gross (New York: Springer) pp 117 - 151

Nakayama K, Silverman G H, 1988 "The aperture problem II. Spatial integration of velocity information along contours" *Vision Research* **28** 747 - 753

Newsome W T, Parré E B, 1988 "A selective impairment of motion perception following lesions of the middle temporal visual area (MT)" *Journal of Neurosciences* **8** 2201 - 2211

Pantle A, Hicks K, 1985 "Using low-level filters to encode spatial displacements of visual stimuli" *Spatial Vision* **1** 69 - 82

Perrone J A, 1990 "Simple technique for optical flow estimation" *Journal of the Optical Society of America A* **7** 264 - 278

Ramachandran V S, Inada V, 1985 "Spatial phase and frequency in motion capture of random-dot patterns" *Spatial Vision* **1** 57 - 67

Ramachandran V S, Cavanagh P, 1987 "Motion capture anisotropy" *Vision Research* **27** 97 - 106

Sereno M E, 1989 "A neural network model of vision motion processing" PhD dissertation, Department of Psychology, Brown University, Providence, RI 02912, USA

Stone L S, Mulligan J B, Watson A B, 1988 "Contrast affects the perceived direction of a moving plaid" *Optical Society of America Annual Meeting* **11** 141

Thompson P, 1982 "Perceived rate of movement depends on contrast" *Vision Research* **22** 377 - 380

Van Santen J P H, Sperling G, 1985 "Elaborated Reichardt detectors" *Journal of the Optical Society of America A* **2** 300 - 321

Wang H T, Mathur B, Koch C, 1989 "Computing optical flow in the primate visual system" *Neural Computation* **1** 92 - 103

Watson A B, Ahumada A J, 1985 "Model of human visual-motion sensing" *Journal of the Optical Society of America A* **2** 322 - 341

Welch L, 1989 "The perception of moving plaids reveals two motion-processing stages" *Nature (London)* **337** 734 - 736

Welch L, Bowne S F, 1989 "Neural rules for combining signals from moving gratings" *Supplement to Investigative Ophthalmology and Visual Science* **30** 75

Wilson H R, 1985 "A model for direction selectivity in threshold motion perception" *Biological Cybernetics* **51** 213 - 222

1095 — 8:30

DETECTING A SINGLE POINT MOVING ON A LINEAR TRAJECTORY AMIDST RANDOMLY MOVING POINTS

Suzanne P. McKee and Scott N. J. Watamaniuk, Smith-Kettlewell Eye Research Institute, San Francisco, CA

A single point moving in apparent motion along a linear trajectory is easily detected when presented against a background of similar points in random apparent motion (Ross, 1988). Since the motion of this single point (the "signal") can be detected even when the spatial-temporal characteristics of the background "noise" are identical from frame-to-frame to that of the signal, the human motion system must integrate the motion signal for many frames within sensors narrowly tuned to particular directions of motion. We measured the detectability of a point moving for 500 msec straight through the center of a ten degree field in one of eight directions, spanning 360 degrees, chosen at random; subjects judged whether the signal point was present or not. Detectability was measured as a function of the increasing density of the noise, an operation that necessarily increased the probability of a mis-match between the signal point and the background points. Surprisingly, subjects could readily detect the signal point ($d' \geq 2.0$) when the probability of a mis-match was as high as 38%, assuming nearest-neighbor matching. Small random perturbations in the straightness of the trajectory ("wobble") had no effect on detectability provided that the directional range of the perturbations did not exceed a bandwidth of 30 degrees. When the motion of the point was broken into small vectors and displayed in random sequence at positions along the trajectory path, detectability decreased significantly. Thus, the ordered sequence, characteristic of natural motion trajectories, appears to enhance the signal within directionally-tuned mechanisms.

AFOSR-89-0035 & Smith-Kettlewell Rachel C. Atkinson Fellowship

McKee, Suzanne P.
AFOSR-89-0035

DO NOT FOLD THIS FORM

1992 ARVO Abstract Form

Numbers on this form refer to Instructions on pages 1-3

3 PRESENTATION PREFERENCE
Please check one (1):
___ (POS) Poster only or Withdraw
___ (NP) Paper or Poster (no preference)
✓ (PA1) Paper #1, Poster #2
___ (POS1) Poster #1, Paper #2

4 SCIENTIFIC SECTION PREFERENCE
See also Section Descriptions on page 14. Please check the one (1) Section best suited to review your proposed presentation.
___ (A) Anatomy/Pathology
___ (B) Biochemistry/Molecular Biology
___ (C) Clinical/Epidemiologic Research
___ (D) Cornea
___ (E) Electrophysiology
___ (F) Eye Movements/Strabismus/Amblyopia
___ (G) Glaucoma
___ (I) Immunology/Microbiology
___ (L) Lens
___ (P) Physiology/Pharmacology
___ (R) Retina
___ (T) Retinal Cell Biology
✓ (V) Visual Psychophysics/Physiological Optics

1 FIRST (PRESENTING) AUTHOR (Must be first author listed in body of abstract.)
☑ 1992 ARVO Member 30845 (Membership Number—Look on Mailing Label)
☐ Non-Member (must have Sponsor—see 2 below)
☐ Applying 1992 ARVO Member (Membership Application Form enclosed).

WATAMANIUK (Last) SCOTT (First) NOLAN JAMES (Middle)
(Department) SMITH-KETTLEWELL (Institution)
2232 WEBSTER ST (Street Address)
SAN FRANCISCO (City) CA, USA (State Country) 94115 (Zip + 4) (Province)
(415) 561-1731 (Office Phone) (415) 561-1610 (FAX) (415) 359-6321 (Home Phone)

2 SPONSOR (Must be a 1992 ARVO member if First Author is not; cannot be First Author or Sponsor on another abstract; and must be a co-author on this abstract.)
(Last) (First) (Middle)
( ) (Office Phone) ( ) (FAX) ( ) (Home Phone)
(Signature of Sponsor) (Membership Number)

5 TOPIC CODE Write the number of one (1) Topic Code that most accurately reflects the content of your proposed presentation. See also page 12.
1 8 2

6 TRAVEL FELLOWSHIP GRANT Put a check in the box if you are applying for a Travel Fellowship Grant. See also pages 15-16.
☐

7 ABSTRACT Do not exceed the borders of the blue box to the right.

8 COPYRIGHT TRANSFER SECTION
Put a check in the box to indicate that all required signatures are on the reverse of this Form.
☑

9 HUMANS/ANIMALS USED IN RESEARCH
Were humans and/or animals used in this research? See also pages 17-19.
☑ Yes ☐ No
IF YES, initials of First Author in box below signifies compliance.
SW

SIMULTANEOUS DIRECTION INFORMATION FROM GLOBAL FLOW AND A LOCAL TRAJECTORY COMPONENT

Scott N. J. Watamaniuk
Smith-Kettlewell Eye Research Institute, San Francisco, CA

Watamaniuk, Sekuler and Williams (1989) measured direction discrimination for global motion produced by dots which randomly chose their movements from a broad range of directions. They showed that thresholds were unchanged whether each dot randomly chose a new direction or chose the same direction each frame as long as the range of directions remained constant. Last year, McKee and Watamaniuk showed that observers can easily detect one dot moving on a fixed trajectory amidst a background of dots that randomly selected their direction of movement each frame from all 360 deg. Does evaluating a global motion preclude the processing of local trajectories?

In the present experiments, 100 dots were randomly assigned directions each frame from a direction distribution with a bandwidth of 120 deg. These dots produced clear global motion in a single direction. Amidst these background dots, one dot moved in a single direction for the duration of the display. The direction of this 'trajectory dot' was in the middle of the distribution range of the background dots. Direction discrimination for both the global motion and local trajectory were measured using the method of single stimuli under a pre-cue and post-cue condition. A tone signalled which motion to judge; in the pre-cue condition the tone was presented 200 msec before the stimulus while in the post-cue condition, the tone was presented immediately AFTER the stimulus. Direction discrimination of either motion was as good in the post-cue condition as in the pre-cue condition. This result suggests that direction information for both global and local motion are processed simultaneously and either can be accessed after the presentation of a stimulus.

Supported by AFSOR 89-0035

McKee, Suzanne P.
AFOSR-89-0035
DO NOT FOLD THIS FORM

1992 ARVO Abstract Form

Numbers on this form refer to Instructions on pages 1-3

1 PRESENTATION PREFERENCE
Please check one (1):
___ (POS) Poster only or Withdraw
___ (NP) Paper or Poster (no preference)
X (PA1) Paper #1, Poster #2
___ (POS1) Poster #1, Paper #2

3 SCIENTIFIC SECTION PREFERENCE
See also Section Descriptions on page 14. Please check the one (1) Section best suited to review your proposed presentation.
___ (A) Anatomy/Pathology
___ (B) Biochemistry/Molecular Biology
___ (C) Clinical/Epidemiologic Research
___ (D) Cornea
___ (E) Electrophysiology
___ (F) Eye Movements/Strabismus/Amblyopia
___ (G) Glaucoma
___ (I) Immunology/Microbiology
___ (L) Lens
___ (P) Physiology/Pharmacology
___ (R) Retina
___ (T) Retinal Cell Biology
X (V) Visual Psychophysics/Physiological Optics

4 TOPIC CODE Write the number of one (1) Topic Code that most accurately reflects the content of your proposed presentation. See also page 12.
182

5 TRAVEL FELLOWSHIP GRANT Put a check in the box if you are applying for a Travel Fellowship Grant. See also pages 15-16.
☐

6 ABSTRACT Do not exceed the borders of the blue box to the right.

7 COPYRIGHT TRANSFER SECTION
Put a check in the box to indicate that all required signatures are on the reverse of this Form.
☒

8 HUMANS/ANIMALS USED IN RESEARCH
Were humans and or animals used in this research? See also pages 17-19
☒ Yes ☐ No
IF YES, initials of First Author in box below signifies compliance
MJB

1 FIRST (PRESENTING) AUTHOR (Must be first author listed in body of abstract.)
☐ 1992 ARVO Member 30639 (Membership Number—Look on Mailing Label)
☐ Non-Member (must have Sponsor—see 2 below)
☐ Applying 1992 ARVO Member (Membership Application Form enclosed)

BRAVO (Last) MARY (First) JAUCH (Middle)
the SMITH-KETTLEWELL EYE RES. INST. (Department / Institution)
2232 WEBSTER ST. (Street Address)
SAN FRANCISCO CA 94115 (City / State / Country / Zip + 4)
(415) 561-1715 (Office Phone) (415) 561-1610 (FAX) (415) 726-9045 (Home Phone)

2 SPONSOR (Must be a 1992 ARVO member if First Author is not; cannot be First Author or Sponsor on another abstract, and must be a co-author on this abstract.)
LAST First Middle
( ) Office Phone ( ) FAX ( ) Home Phone
Signature of Sponsor Membership Number

SPEED SEGREGATION AND TRANSPARENCY IN RANDOM DOT DISPLAYS

Mary J. Bravo and Scott N. J. Watamaniuk
The Smith-Kettlewell Eye Research Institute, San Francisco, CA

The visual system can use local speed information to determine whether one surface or two transparent surfaces are visible. Given the number of surfaces, the visual system organizes speed information to generate the signals used for speed discrimination. The following examples illustrate this point.

(1) When half the dots in a random dot cinematogram move upward at a slow speed (about 6 deg/sec) and half move upward at a fast speed (about 2[illegible] deg/sec), two transparent surfaces are seen. Discrimination of small changes in the speed of one set of dots is unaffected by the presence of the other dots.

(2) When the dots alternate *synchronously* between the two speeds so that any instant only one speed is present, then one surface is seen. For all alternation rates tested, discrimination of either speed is greatly impaired.

(3) When the dots alternate *asynchronously* between the two speeds, so that any instant both speeds are present, then two transparent surfaces that *twinkle* are seen. (Note that since each dot has a continuous upward trajectory, the percept of twinkle must arise from the segregation of the two speeds making up the trajectory.) For all but the fastest alternation rates, discrimination of one speed is unimpaired by the presence of the other speed.

In addition to alternation rate, we find that the spatial properties of the display influence segregation and transparency. These results are interpreted in terms of the rules governing the segregation of surfaces.

Supported by NEI F32-EY06155 and AFSOR 89-0035.

ONE-PATH MODEL FOR CONTRAST-INDEPENDENT PERCEPTION OF FOURIER AND NON-FOURIER MOTIONS Norberto M. Grzywacz, Smith-Kettlewell Institute, 2232 Webster Street, San Francisco, CA 94115.

A class of motion perception models is based on motion-energy mechanisms, which are spatio-temporally oriented filters (Adelson and Bergen, 1985). A population code of such filters may provide the basis for visual velocity computation (e.g., Grzywacz and Yuille, 1989). A problem with these filters is that they cannot account for percepts in which the relevant Fourier components do not move with the perceived motion. Moreover, against human psychophysical evidence, the performance of motion-energy mechanisms improves with contrast. To solve these problems, we extended the Grzywacz and Yuille model by adding a rectified band-pass filter in front of the motion-energy filters (Chubb and Sperling, 1988). The extension also postulated that the system's limiting noise is multiplicative and at the stage combining the filters' outputs.

Mathematical and computational calculations show that this extended model accounts for motion perception of stimuli for which Fourier analysis yields no systematic motion components (Chubb and Sperling, 1988). Furthermore, the extended model is consistent with a large set of "beat-pattern" motion phenomena (e.g., Derrington and Baddock, 1985). The model may also explain the invariance of the speed Weber fraction with contrast (e.g., McKee et al., 1986). Finally, it can be shown that the model can deal with motion transparency (Smith and Grzywacz, 1992). We conclude that two parallel motion pathways (e.g., Wilson, 1991) are not necessary to account for Fourier and Non-Fourier motions.

Supported by AFOSR-89-0035

Poster

PARTICIPANT INFORMATION: POSTER
(To be printed in convention program)

Complete this page, and the PROPOSAL COVER SHEET, and send ***six printed copies*** of both forms to APS, 1010 Vermont Ave, NW, Suite 1100, Washington, DC 20005-4907. **Required:** Each submitter must send either a PC-compatible diskette (3.5" or 5.25") or Macintosh-compatible diskette (3.5") with a file that contains the title, each author's name and affiliation, abstract, and two-digit subject index code(s). The IBM-compatible diskette file should be in ASCII/DOS format and named with the ".TXT" file extension, commonly used with ASCII files. The Apple/Macintosh diskette file should be a MacWrite or text-only format file. (Make sure you provide the ASCII or text file and not an idiosyncratic file created by your particular word processor.) **The diskette label should indicate IBM or Apple/Macintosh compatibility and list: (1) senior author's name, (2) session format, (e.g., poster, symposium), and (3) key word(s) from the title.**

1. Title (12 words or fewer):
   Transparency Influences Speed Discrimination in Random Dot Displays

2. First author's name, affiliation, and mailing address:
   Scott N.J. Watamaniuk, Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115

3. Second author's name and affiliation:
   Mary Bravo, Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115

4. Third author's name and affiliation:

5. Fourth author's name and affiliation:

6. List additional authors on another page if needed.

7. Abstract (50 words or fewer):

   An observer's ability to discriminate one component speed of a display composed of dots moving upward at two very different speeds, depends on the way the speeds are distributed in space and time. Good speed discrimination is closely linked to the perception of transparency.

8. List Subject Index Code(s) (Use number after psychology speciality(ies), listed on Cover Sheet.):
   Subject Index Codes: 58, 52, 23

9. Which authors are APS members? (At least one author must be an APS member. If all authors are students, list the sponsoring APS member here and include his/her signature.)
   Scott N. J. Watamaniuk

Proposal Summary (Poster)

The visual system can use local speed information to determine whether one surface or two transparent surfaces are visible. For example, we observed that two sets of spatially intermingled dots, moving in the same direction but at very different speeds, segregate into two transparent planes sliding over each other. Moreover, observers could discriminate the speed of either the fast or slow component as precisely as if that component were presented alone - the presence of the other speed did not influence speed discrimination. Our research examined the stimulus parameters that produce transparency in random dot displays.

Our stimuli were random dot cinematograms in which dots moved in the same direction (upwards) at a slow speed (about 6 deg/sec) and at a fast speed (about 21 deg/sec). Welch and Bowne (1990) found that for plaid stimuli composed of two differently oriented moving gratings, good speed discrimination of one of the component gratings was only possible when the components segregated. Therefore, we used speed discrimination of the component speeds in our display as an index of segregation. The observer was told in advance which component (fast or slow) he/she was to discriminate. Stimuli were presented using a variant of the method of constant stimuli, known as the single-stimulus method (McKee & Welch, 1985). In each trial, one cinematogram was presented and the observer's task was to determine whether the designated component speed, in the current stimulus, moved faster or slower than the mean speed of the stimulus set.

We tested several stimulus conditions. In the baseline condition, cited above, two sets of spatially intermingled dots moving at a different speeds were presented. In the other two conditions, dot speeds alternated between the two component speeds. The alternation of the dots' speeds were either **synchronous**, so that at any instant only one speed was present, or **asynchronous** so that both speeds were present in the display simultaneously. We measured speed discrimination for both conditions, at four different alternation rates: 1, 2, 4, or 8 frames/alternation.

We found that when dots alternated their speeds synchronously, only one surface was seen and speed discrimination of the component speed was very poor. This result was obtained for all four alternation rates. However, when the dots alternated their speeds asynchronously at 4 or 8 frames/alternation, two twinkling transparent surfaces were seen (since each dot had a continuous upward trajectory, the percept of twinkle must have arisen from the segregation of the two speeds making up the trajectory). At these two alternation rates, speed discrimination for the components was as precise as in the baseline condition. Note that individual dots changed their speeds in an identical manner in both the synchronous and asynchronous conditions; these two condition are not discriminable given the behavior of a single dot. These conditions differed only in the pattern of speed alternation across dots.

We conclude that segmentation from speed information requires specific spatial-temporal relationships between the two different speeds (e.g., both speeds must be simultaneously present). In addition, when segmentation occurs, the visual system organizes speed information accordingly to generate the signals used for speed discrimination.

DO NOT FOLD THIS FORM

1992 ARVO Abstract Form

Numbers on this form refer to Instructions on pages 1-3

3 PRESENTATION PREFERENCE
Please check one (1):
___ (POS) Poster only or Withdraw
___ (NP) Paper or Poster (no preference)
X (PA1) Paper #1, Poster #2
___ (POS1) Poster #1, Paper #2

4 SCIENTIFIC SECTION PREFERENCE
See also Section Descriptions on page 14. Please check the one (1) Section best suited to review your proposed presentation.
___ (A) Anatomy/Pathology
___ (B) Biochemistry/Molecular Biology
___ (C) Clinical/Epidemiologic Research
___ (D) Cornea
___ (E) Electrophysiology
___ (F) Eye Movements/Strabismus/Amblyopia
___ (G) Glaucoma
___ (I) Immunology/Microbiology
___ (L) Lens
___ (P) Physiology/Pharmacology
___ (R) Retina
___ (T) Retinal Cell Biology
X (V) Visual Psychophysics/Physiological Optics

1 FIRST (PRESENTING) AUTHOR (Must be first author listed in body of abstract.)

☒ 1992 ARVO Member 27406 (Membership Number—Look on Mailing Label)
☐ Non-Member (must have Sponsor—see 2 below)
☐ Applying 1992 ARVO Member (Membership Application Form enclosed).

Grzywacz (LAST) Norberto (First) Mauricio (Middle)
(Department) Smith-Kettlewell Institute (Institution)
2232 Webster Street (Street Address)
San Francisco (City) CA (State/Country) 94115 (Zip + 4) (Province)
(415) 561-1795 (Office Phone) (415) 561-1610 (FAX) (415) 332-6428 (Home Phone)

2 SPONSOR (Must be a 1992 ARVO member if First Author is not; cannot be First Author or Sponsor on another abstract; and must be a co-author on this abstract.)

( ) (LAST) ( ) (First) ( ) (Middle)
( ) (Office Phone) ( ) (FAX) ( ) (Home Phone)
(Signature of Sponsor) (Membership Number)

5 TOPIC CODE Write the number of one (1) Topic Code that most accurately reflects the content of your proposed presentation. See also page 12.
1 0 a

6 TRAVEL FELLOWSHIP GRANT Put a check in the box if you are applying for a Travel Fellowship Grant. See also pages 15-16.
☐

7 ABSTRACT Do not exceed the borders of the blue box to the right.

8 COPYRIGHT TRANSFER SECTION
Put a check in the box to indicate that all required signatures are on the reverse of this Form.
☒

9 HUMANS/ANIMALS USED IN RESEARCH
Were humans and/or animals used in this research? See also pages 17-19.
☐ Yes ☒ No
IF YES, initials of First Author in box below signifies compliance.
☐

ONE-PATH MODEL FOR CONTRAST-INDEPENDENT PERCEPTION OF FOURIER AND NON-FOURIER MOTIONS

Norberto M. Grzywacz, Smith-Kettlewell Institute, 2232 Webster Street, San Francisco, CA 94115.

A class of motion perception models is based on motion-energy mechanisms, which are spatio-temporally oriented filters (Adelson and Bergen, 1985). A population code of such filters may provide the basis for visual velocity computation (e.g., Grzywacz and Yuille, 1989). A problem with these filters is that they cannot account for percepts in which the relevant Fourier components do not move with the perceived motion. Moreover, against human psychophysical evidence, the performance of motion-energy mechanisms improves with contrast. To solve these problems, we extended the Grzywacz and Yuille model by adding a rectified band-pass filter in front of the motion-energy filters (Chubb and Sperling, 1988). The extension also postulated that the system's limiting noise is multiplicative and at the stage combining the filters' outputs.

Mathematical and computational calculations show that this extended model accounts for motion perception of stimuli for which Fourier analysis yields no systematic motion components (Chubb and Sperling, 1988). Furthermore, the extended model is consistent with a large set of "beat-pattern" motion phenomena (e.g., Derrington and Baddock, 1985). The model may also explain the invariance of the speed Weber fraction with contrast (e.g., McKee et al., 1986). Finally, it can be shown that the model can deal with psychophysical data on motion transparency and coherence (Smith, Grzywacz, and Hildreth, 1992). We conclude that two parallel motion pathways (e.g., Wilson, 1991) are not necessary to account for Fourier and Non-Fourier motions.

Supported by AFOSR-89-0035